DISCOURSE COHESION AND PHONETICS IN TURKISH SIGN LANGUAGE (TİD): AN EXPERIMENTAL AND COMPUTATIONAL APPROACH

ONUR KELEŞ

BOĞAZİÇİ UNIVERSITY

2024

DISCOURSE COHESION AND PHONETICS IN TURKISH SIGN LANGUAGE (TİD): AN EXPERIMENTAL AND COMPUTATIONAL APPROACH

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Linguistics

by

Onur Keleş

Boğaziçi University

2024

Discourse Cohesion and Phonetics in Turkish Sign Language (TİD): An Experimental and Computational Approach

The thesis of Onur Keleş

has been approved by:

Assist. Prof. Kadir Gökgöz (Thesis Advisor)	______________________
Assist. Prof. Nazik Dinçtopal Deniz (Thesis Co-Advisor)	______________________
Assoc. Prof. Didar Akar	______________________
Assist. Prof. Stefano Canalis	______________________
Prof. Albert Ali Salah (External Member)	______________________

December 2024

DECLARATION OF ORIGINALITY

I, Onur Keleş, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature: ....

Date: ....

ABSTRACT

Discourse Cohesion and Phonetics in Turkish Sign Language (TİD): An Experimental and Computational Approach

Theories of linguistic efficiency, such as Zipf's (1949) law, claim that languages reduce effort by favoring simpler or more economical forms whenever possible. Although this claim has widely been tested and confirmed in spoken languages (e.g., Givon, 1983; Gundel, Hedberg, & Zacharski, 1993), it has not been addressed to the same extent in sign languages. This thesis investigates how such theories would be realized in Turkish Sign Language (TİD) through discourse cohesion and phonetic reduction in the narratives of TİD. With a story-retelling production experiment, discourse cohesion has been analyzed using a quantized measure of accessibility adapted from Ariel's (1990) framework. In particular, I examine referring expression forms (e.g., nominal versus verbal) that native and late-signing deaf adult TİD signers employ in narratives. The articulatory phonetic aspect of narratives has been analyzed using MediaPipe, an open-source computer vision tool. The results of the discourse cohesion experiment indicated a strong relationship between the cognitive accessibility scores of referents and the discourse context (e.g., first mention, maintenance, re-introduction), type of the referring expression, and age of acquisition in TİD, which aligned well with previous research on spoken language and other sign languages. The results of the computational phonetic analysis showed the forms that TİD signers used underwent phonetic reduction as the cognitive accessibility of a referent increased. They had shorter duration, smaller hand movement, and narrower signing space. Age of acquisition or delayed first acquisition did not significantly affect these measures except for duration, in which native signers retold the events faster than late signers. In conclusion, this thesis reconciles discourse cohesion and phonetic reduction within the framework of linguistic efficiency by offering a novel methodology to analyze signed discourse using both an experimental and computational approach.

ÖZET

Türk İşaret Dilinde Söylem Bütünlüğü ve Hareket Analizi: Deneysel ve Hesaplamalı Bir Yaklaşım

Dil verimlilik teorileri bireylerin mümkün olduğunca daha basit veya daha ekonomik formları tercih ettiğini ve harcadıkları çabayı azalttığını ileri sürer. Bu tezde, bu verimlilik eğilimi, Türk İşaret Dili'ndeki (TİD) söylem bütünlüğü ve hesaplamalı hareket analizi üzerinden incelenmiştir. Söylem bütünlüğü, Ariel'in (1990) çalışmasından uyarlanan bir erişilebilirlik ölçüsü kullanılarak analiz edilmiştir. Özellikle, bir hikaye anlatma deneyi ile TİD'i anadil olarak ve geç edinmiş sağır yetişkin bireylerin anlatılarda kullandıkları gönderimsel ifade dizgileri incelenmiştir. Anlatılarda kullanılan hareketlerin hesaplamalı kinematik analizi, bir poz tahminleme programı (MediaPipe) kullanılarak analiz edilmiştir. Sonuç olarak, hesaplanan erişilebilirlik puanlarının gönderim formunun (örn. isim veya eylemsel gönderim), söylem bağlamı (ilk bahis, sürdürme, yeniden bahsetme) ve dil edinim grubuyla ilişkili olarak değiştiği gözlemlenmiştir. Hesaplamalı analiz sonuçları ise kinematik değerlerin söylem bağlamı ve bir gönderim erişilebilirliğine göre değişiklik gösterdiğini ortaya koymaktadır. TİD kullanan sağır bireyler, bir gönderimin bilişsel erişilebilirliği arttıkça o işareti daha kısa süreli, daha küçük el hareketleriyle ve daha dar bir işaretleme alanında yapmışlardır. Dil edinim durumu ise anlatı süresi dışında bu kinematik değerler üzerinde anlamlı bir etki yaratmamıştır. Anadil olarak TİD edinmiş sağırlar anlatıları geç edinim grubuna göre daha hızlı anlatmışlardır. Sonuç olarak, bu tezde, TİD anlatılarında bir söylem bütünlüğü ve hareket analiziyle sağır yetişkinlerin kullandıkları formu gönderimlerinin erişebilirliğine ve bağlamına göre değiştirdikleri ve dili daha ekonomik kullandıkları görülmüştür. Bu da iletişim kanalından bağımsız olarak dillerin evrensel verimlilik ve düşük çaba ilkelerini sağladığı savını desteklemektedir.

ACKNOWLEDGEMENTS

First and foremost, I extend my heartfelt thanks to my advisor, Dr. Kadir Gökgöz. I was just a second-year undergraduate student when I nervously walked into his office to share my research ideas about TİD. To my surprise, he not only took me seriously but also offered me a role in his research project, providing funding to conduct my experiment. Despite my lack of experience at the time, Kadir Hoca always believed in me and pushed me to aim higher than I ever thought possible. He encouraged me to submit my results to top-tier conferences and SSCI journals, even though I faced rejections initially. Thanks to his never-ending support, superb guidance, and incredible mentorship, I had the privilege of publishing or presenting in such venues under his supervision, doing my first-ever TA-ship in his LING 101 class, and receiving his unwavering support during my grad school and RA applications. I was also very grateful for the opportunity to work as an assistant and then an instructor at the Turkish Language and Culture Program. These all shaped me into the person and researcher I am today.

Beyond academic growth, Kadir Hoca has always helped me whenever I face doubts about myself or feel uncertain about the future. He listened with patience and understanding and was there for me. As a side note, he is among the funniest and kindest people I know!

I must wholeheartedly thank my co-advisor, Dr. Nazik Dinçtopal Deniz, who kindly agreed to supervise this thesis on the condition that I also do an eye-tracking study in TİD with her, which was an immediate yes! From the moment I stepped into Nazik Hoca's class in undergrad, which was phenomenally taught, I knew I wanted to work with her. I was lucky enough to join her psycholinguistics lab in my senior year, in which I learned a lot about language processing and experimental design. I am also very grateful for her meticulous feedback on my thesis, which I believe is in much better shape than its earlier form.

I also need to thank the committee members. I am thankful to Dr. Didar Akar, from whom I took my first LING 101 class in Fall 2017 and I feel very honored to do so. Didar Hoca's comments and suggestions were incredibly useful. I thank Dr. Stefano Canalis for being kind and understanding every time I wanted to meet for feedback on my thesis. I am also grateful for the classes I took with him, both as an undergraduate and as a graduate student. I also thank Dr. Albert Ali Salah for graciously agreeing to be on the committee. I am especially thankful for the time he dedicated to inviting me to Utrecht University during my visit and for his thoughtful feedback on my thesis draft.

There are a few institutions I would like to thank for the financial support I received during my masters. The Scientific and Technological Research Council of Turkey (TÜBİTAK) provided support through the 2210-A National Scholarship Programme for Master's Students, and the Dil Uygulama ve Araştırma Merkezi generously funded my conference travels. In addition, I must thank Elvan Tamyürek Özparlak and Buket Ela Demirel for their assistance in data collection and annotation, and of course, all the participants.

The Department of Linguistics has been my academic home for over three years, during which I have had the privilege of serving as a TA for half of that time. Dr. Balkız Öztürk is the role model every linguist aspires to be. I am nothing but honored to have met her and even become her TA. Honestly, one of my proudest achievements is being able to say, 'My syntax teacher was Balkız Hoca', and that statement can speak for itself. She also supported me in any way possible -not just because I'm a Virgo- and always managed to calm my work-related anxiety when things went awry.

I am also thankful to Dr. Ümit Atlamaz. As a senior undergrad, I attended his computational linguistics class, which turned out to be a pivotal moment in shaping my academic interests and career trajectory. Moreover, assisting Ümit Hoca's classes in the last two years has definitely been an eye-opening experience for me and helped me become a better TA. My acknowledgements extend to Dr. Ömer Demirok, from whom I was lucky enough to take grad syntax and morphology. Ömer Hoca excelled in simplifying complex concepts and made literally everything more engaging thanks to his wonderful personality, so at some point, an experimental linguist like me even thought of doing nanosyntax for a career. My only regret is that I wish I could have TA'ed for him much earlier since I am learning so much.

Without doubt, I must thank Dr. Metin Bağrıaçık, Dr. Elena Guerzoni, Dr. Cem Bozşahin, Dr. Mine Nakipoğlu, Dr. Pavel Logacev, Dr. Sumru Özsoy, and our departmental secretary Gülsüm Abay. Metin Hoca is the kindest person ever and having the TA office right across his office makes my mornings better and definitely alleviates my morning depression. I thank him for always having some time whenever I need consultation about anything. Without Elena Hoca's challenging but fun assignments, and top-notch lectures in graduate semantics, I would have already given up on lambda calculus by now. I also need to acknowledge our fun conversations in the corridor on anything ranging from Halloween costumes to cats. Of course, I must express my gratitude to Cem Hoca, who is one of the kindest and smartest people I know. I owe a lot to him for this thesis because not only did he spend hours listening to my research results but also he gave me very valuable feedback and lots to think about. Mine Hoca always fueled my passion for pursuing cognitive linguistics, and TA'ing for her for the first time made me truly appreciate how incredible her lectures are. Pavel Hoca taught me all I know about Bayesian statistics and R. I owe him a lot for all the discussions. Sumru Hoca is the one who introduced me to the Deaf community and structure of sign languages. I will always be grateful to be her student. It is a true honor. Finally, Gülsüm Hanım deserves an acknowledgement. I cannot remember the number of times I asked for help with a problem and she always found a way to solve it with a smiling face.

I was also lucky to visit Max Planck Institute for Psycholinguistics and met great people there. I thank Dr. Aslı Özyürek for many insightful and thought-provoking conversations about my potential thesis topic as well as all the MPIers I found the chance to talk with. Another special thanks go to Dr. Didem Gündoğdu, who introduced me to the world of computational social science. Didem Hoca's course and her impeccable mentorship paved the path to my thesis. Finally, I thank Dr. Bahtiyar Makaroğlu, Dr. Gerardo Ortega, Dr. Ann Senghas, Dr. Roland Pfau, Dr. Pamela Perniss, Dr. Esam Ghaleb, and Dr. Julia Krebs for meeting with me and giving me invaluable feedback on early drafts of my thesis.

I want to express my gratitude to friends: Mehmet Can Dadan, Metehan Eryılmaz, and Ömer Tabak for being awesome, kind, and helpful office mates. Assisting next to you has been a true pleasure. Muhammed İleri and Furkan Dikmen who taught me pretty much everything I know about this job. Although we spent little time as colleagues, I will always cherish our conversations. And to the best cohort ever: Baran Günay, Kaan Bayar, Nagihan Lüle, Rüveyda Sahyar, Sena Çiftçi, Zeynep Erdemir, and many others I cannot name, for creating such a positive atmosphere in the department and being such nice friends. As a side note, thanks Baran for becoming my *edi* when we fine-tuned *büdü*. Also, shout out to everyone who contributed to the departmental potlucks. Highlights include *yaprak sarma*, tiramisu, and spinach labneh pastry. My thanks extend to Ceren Oksal, who became my thesis and hackathon buddy, for her amazing friendship and support. Demet Kayabaşı for all the sign language talks we had and help in collecting the data for this thesis, Aslı Kuzgun for all the fun conversations and gossips, Karahan Şahin, my go-to person whenever I had a Python or CV question, Hande Sevgi and Bergül Soykan for being my Boston-buddies, Büşra Marşan for being such a nice person, Furkan Atmaca for teaching me how to code in R and preparing this LaTeX thesis template, Utku Türk for countless conversations on methodology and R. I also thank my data science buddies Ömer Bayraklı, who has also become a brother and a close friend to me, Berat Doğan, Mert Gökyar and the Compec community for the support.

Finally, I thank my family, who brought me here to day. Without their patience and support, I would have not been here. Most importantly, my *dearest* wife, Zeynep İrem Bayraklı Keleş. My life began 8 years ago when I first met you. The rest is story. This thesis is dedicated to you. Thank you for being my partner in crime, my best friend, soulmate and my love. Also, thank you for letting me write this thesis for hours and skipping some house chores. I'll make it up to you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Effective communication in any language requires a balance between clarity and brevity. Interlocutors try to make their messages understandable without unnecessary redundancy or superfluous use of form. This aim is justified with certain cognitive principles that encourage efficiency, which suggests that we naturally favor less effortful or more economical forms of expression. This simple idea can be extended to discourses in which speakers refer to entities. There, the drive for the economy of form is further shaped by theories of discourse cohesion. Such theories explain how forms are chosen based on how accessible or cognitively active a referent is assumed to be in the addressee's mind at a given point in the discourse. Suppose a referent is highly accessible (e.g., it has been recently mentioned or remains prominent in the conversation). In that case, we choose simpler or reduced forms encapsulating the least redundant information, such as pronouns or null forms, which refer to cases where we omit the referent anaphor (Ariel, 1990; Givon, 1983; Gundel, Hedberg, & Zacharski, 1993). In contrast, if the referent is less accessible (e.g., it was mentioned much earlier or is less prominent), we use a more explicit form, like a full-noun phrase, to avoid possible ambiguities. For example, given a short discourse like in (1), the forms "Eddie", "he", and "∅" (zero anaphora) all refer to the same individual.

(1) So Eddie turned around. He said, "youse got a problem?" "Yeah, we want you," they say. So—∅ walked right up to them, and they just pulled him down (Williams, 1988, p. 343, ex. 7)

If the contextual and linguistic information about the referent is abundantly available to the recipient, less of a quantity of form is used to satisfy a certain threshold to refer to an entity effectively. If not, forms with more phonetic quantity must be employed to reach the accessibility threshold (Czubek, 2017). This has the following implication: on the one hand, when producing long discourses like narratives, languages must be expressive enough to avoid ambiguity and communicate the message successfully. On the other hand, they must minimize the biomechanical articulatory effort by means of phonetic reduction. As such, discourse cohesion follows from different theoretical frameworks, including the principle of least effort (Zipf, 1949), the Griciean maxim of quantity (Grice, 1975), relevance theory (Sperber & Wilson, 1986), and cognitive accessibility hierarchies (one example is Ariel, 1990), all of which will be elaborated in Chapter 2.

The principle of referencing and economy of form is not peculiar to only one communication modality. Sign languages are the natural languages used by Deaf[1] communities worldwide. These are independent languages that do not originate from the surrounding spoken languages. Instead, they develop naturally when Deaf individuals come together to form a community. As per building a cohesive discourse in sign languages, these dynamics unfold within a visual-spatial modality, where the physical constraints and affordances of the signing body influence the economy of form and accessibility of referents. However, we know from existing literature that the same principles apply to signed discourse (e.g., Ferrara et al., 2022; Frederiksen & Mayberry, 2016; Swabey, 2002). Signers use more attenuated or reduced forms for familiar or accessible referents, and they use fuller forms for less accessible referents.

It has also been reported that building a successful, cohesive discourse is challenging for second-language learners since it demands the simultaneous integration of syntactic and discursive properties (Sorace, 2011; Sorace & Filiaci, 2006). This often leads to using more form than necessary, or in other words, being overexplicit in referencing. In sign languages, the existing research on the age of acquisition has expanded, addressing the effects of delayed first language exposure, which is common among deaf children born to hearing parents. Extensive work has emerged showing that this delay can affect language production and processing in adulthood, particularly for complex morphosyntactic structures (Boudreault & Mayberry, 2006; Kayabaşı & Gökgöz, 2023; Sevgi & Gökgöz, 2023). Nevertheless, a lot less is known about the effects of delayed first language acquisition on discourse cohesion or reference tracking. This thesis addresses this gap by providing insights

[1] I use Deaf with a title case to refer to the culture and community.

into the retellings of both native and late-signing deaf adults through a production experiment.

Given the recent developments in computer vision, new methodologies can now quantify various phonetic measures in sign languages. Mainly, open-source pose estimation tools, such as OpenPose (Cao, Simon, Wei, & Sheikh, 2017) and MediaPipe (Lugaresi et al., 2019), have been applied to sign language research (see Börstell, 2023), and this enables the measurement of visual-kinematic features like speed, movement, and velocity in the visual-spatial modality. This is possible thanks to calculations like landmark detection and tracking for human bodies in 2D videos, estimating the poses of joints on the x, y, and z axes. As per reference tracking, these computational metrics can now be utilized to detect phonetic reduction in signed discourse and directly examine the relationship between referential accessibility and phonetic marking on the form. That is, the extent of reduction can be assessed to determine if more accessible referents are being produced with less effort, consistent with proposed accessibility hierarchies. A few studies reported reduction in phonetic form for predictable or repeated references for sign languages (Hoetjes, Krahmer, & Swerts, 2014; Stamp, Dachkovsky, Hel-Or, Cohe, & Sandler, 2024).

This thesis investigates how principles of efficiency are realized in the narratives of Turkish Sign Language, focusing on the discourse cohesion and phonetic strategies that deaf adult native and late signers use to track referents. With this research, I aim to contribute to a broader understanding of the economy of language, referential accessibility, and the impact of age of acquisition on both pragmatic and phonetic aspects of language use. Ariel's 1990 accessibility hierarchy, was adopted in this research. As such, the primary question of interest was how accessibility might shape form selection (e.g., nominal versus verbal) and a form's kinematic or phonetic properties. A story-telling production experiment was conducted to elicit retellings, and an open-source computer vision tool, MediaPipe, was employed to examine the relationship between discourse cohesion and phonetics. By examining these factors, this thesis investigated not only the pragmatic and phonetic efficiency with which native and late TİD signers maintain discourse cohesion but also the extent of reduction and variation in quantized kinematic metrics across referents.

The layout of the thesis is as follows: Chapter 2 provides theoretical background information on discourse cohesion and the prominent referring expressions used in sign languages, as well as an overview of the age of acquisition effects among signers. The chapter concludes with a review of phonetic reduction and computer vision applications, connecting these topics to the broader goals of the thesis.

Chapter 3 outlines the research questions and methodology, detailing the production experiment and the computer vision analysis, including participants, task stimuli, procedures, data annotation, and accessibility score calculation.

Chapter 4 examines discourse cohesion analysis in TİD narratives. The chapter first explores discourse cohesion, presenting the data analysis procedure and the results on how discourse context and group differences impact accessibility scores and the effects of referring expression types.

Chapter 5 presents the phonetic analysis, starting with the description of landmark extraction and data preprocessing, and then outlines the data analysis procedure and the results for both a lexical-(item-)level and an utterance-level phonetic analysis.

Chapter 6 presents a general discussion of the findings for discourse cohesion and phonetic analyses.

Finally, Chapter 7 concludes the thesis, summarizes the findings and their implications, discusses the limitations of the thesis, and raises directions for future research.

I share all the data and the R and Python scripts used for statistical analysis and computer vision on this online repository:
https://github.com/kelesonur/MA_Thesis

CHAPTER 2

LITERATURE REVIEW

This chapter provides background information on theories of linguistic efficiency and referential accessibility, tools for establishing cohesion in signed discourse, and age of acquisition effects among deaf signers. It also overviews phonetic reduction and computer vision and concludes by listing the motivations for the present study.

A French philosopher, Guillaume Ferrero (1894), made the following observation: we often take shortcuts or use simple heuristics that try to minimize biomechanical effort when performing different forms of physical and mental activities. Almost half a decade later, the American philologist George Kinsley Zipf (1949) posited, by building on Ferrero's idea, the notion of the Principle of Least Effort, also known as Zipf's Law, which maintains that human behavior, including linguistic choices, is guided by a natural inclination to minimize effort while maximizing efficiency. In the context of linguistics, this principle manifests in the tendency of speakers or signers to favor shorter, more economical forms of expression, mainly when referring to frequently mentioned entities. In his work, Zipf analyzed James Joyce's Ulysses and saw that high-frequency words (e.g., give) tend to be shorter and more phonetically reduced than low-frequency words (e.g., ineluctable), reflecting an equilibrium between the interlocutor's effort and the communicative needs of the listener. As the latter work suggests (Ariel, 1990; Givon, 1983; Gundel, Hedberg, & Zacharski, 1993, to name a few), this phenomenon is also apparently evident in building a cohesive discourse and in the tracking of discourse referents, where communicators utilize reduced forms like pronouns or zero anaphora for referents that are highly accessible or previously established in the discourse, thereby economizing on articulatory effort without sacrificing comprehension.

Narratives, which are the subject of this thesis, were defined by Ryan (2003, p. 1) and this thesis will adopt this particular definition:

> "Narrative is defined as a mental image, or cognitive construct, which can be activated by various types of signs. This image consists of a world (setting) populated by intelligent agents (characters). These agents participate in actions and happenings (events, plot), which cause global changes in the narrative world. Narrative is thus a mental representation of causally connected states and events which captures a segment in the history of a world and of its members."

The notion of least effort can play an essential role in narrative production, too. While speakers or signers aim to conserve effort with phonetic reduction, they must also ensure that their tracked referent remains unambiguous.

In narratives, this becomes particularly crucial as the complexity of the tracking task increases with an increasing number of referents. This trade-off between effort and clarity is essential for effective communication, as more economical expressions may lead to clarity. In contrast, excessively detailed expressions can burden both the speaker and listener. Zipf's work thus provides a foundational understanding of how linguistic economy influences referential choices and contributes to discourse cohesion[2]. This idea is consistent with Paul Grice's (1975) Maxim of Quantity, which suggests that language users should give the ideal quantity of information, not too much, but not too little, just enough to be informative. Grice's principle can be subdivided into two tenants, as given in (1).

(1) T1: Make your contribution as informative as is required. T2: Do not make your contribution more informative than is required.

(Grice, 1975, p. 45)

Both approaches indicate a balance between communicative efficacy and language economy. Zipf's principle addresses them at a phonological and lexical level, and Grice's Maxim of Quantity functions at the pragmatic level, guiding speakers to avoid excessive detail while ensuring clarity.

The Relevance Theory, proposed by Sperber and Wilson (1986), further complements these ideas by claiming that the listener operates under Relevance when

[2] I use discourse cohesion interchangeably with reference tracking.

processing incoming utterances. According to this assumption, the addressee generates many contextual implications and collects relevant information using principles such as Maximal Relevance and Optimal Relevance. Sperber and Wilson reduce Grice's maxims to a single notion that is relevant (Johnson-Laird, 1988).

To summarize, all these three frameworks converge on the idea that communication is driven by an overarching principle of cognitive economy, where individuals try to optimize both the form and content of their messages. The following subsections present how the cognitive economy shapes discourse cohesion across spoken and sign languages through form choice and phonetic reduction, given Ariel's concept of referential accessibility.

2.1 Accounts of referential accessibility and discourse cohesion

Extending Zipf's and Grice's idea of the economy of language to discourse cohesion, multiple accounts have been proposed for referential accessibility and how it relates to the choice of form. In this subsection, different linguistic theories and frameworks developed for spoken or written languages that explain how we choose referential forms based on accessibility and discourse cohesion will be explored. Some of these approaches are rooted in the concept of givenness, such as the models proposed by Chafe (1976) and Prince (1978), which focus on shared knowledge and the cognitive status of information—others, such as Givón's functional-typological perspective, center on topic continuity. Moreover, Ariel's accessibility hierarchy will be presented, a more flexible model that combines ideas from givenness and discourse processing to explain these choices of form. Together, these theories show how cognitive, pragmatic, and discourse-related factors influence the way we refer to entities in language.

Givenness, rooted in the Praguian school of linguistics, plays a central role in describing how speakers structure their messages, as it determines the extent to which information is already known or assumed to be known by the listener. Scholars such as Chafe (1976) and Prince (1978) exemplify how the givenness of a referent influences linguistic choice. Chafe relates givenness to consciousness and memory, suggesting that given information remains active in the addressee's mind. That is, given information is still easily accessible because it is within the listener's immediate focus or short-term memory. When this information remains 'active', the speaker can use more economical or reduced forms, such as pronouns or ellipses, to refer to it. However, when information is not given or is no longer at the forefront of the addressee's mind, perhaps because it has not been recently mentioned or because other topics have taken precedence, speakers employ more explicit cues, like full noun phrases, reactivate information and re-establish its relevance in the conversation.

Gundel et al. (1993) also address referential forms through the idea of givenness. Her givenness hierarchy offers a cognitive-pragmatic framework for spoken English that links the referring forms to referential accessibility or, simply put, the cognitive status of entities in the addressee's mind. The hierarchy focuses on six cognitive statuses, ranging from 'in focus' to 'type identifiable,' each associated with specific linguistic forms. For example, "in focus' referents are typically referred to using pronouns, while those that are 'uniquely identifiable' might be referred to with full noun phrases or descriptions such as 'the dog' or 'the car,' which signal that the listener can identify the entity based on shared knowledge or prior mention.

Givón (1983) explores discourse cohesion more from a functional-typological perspective of topic continuity, which is assumed to be unmarked in human languages. Givón identifies factors such as topicality, agency, and continuity that influence the coding of referents, arguing that languages employ a range of devices—like pronoun use, topicalization, and verb agreement—to signal the continuity or shift of topics within discourse. Especially, continuity may imply easier processing due to shared topics (Wang, 2022). In his work, Givón demonstrates that high-topic continuity correlates with reduced or attenuated referential forms, whereas low-topic continuity necessitates more explicit and elaborated expressions and proposes a scale of topic continuity. This scale includes grammatical tools used for topics ranging from most continuous (e.g., zero anaphora) to most discontinuous (e.g., referential indefinite noun phrases).

Built upon Sperber and Wilson's (1986) notion of accessibility, Ariel's (1990) accessibility hierarchy extends these ideas but offers a more flexible and dynamic model for understanding referential choices. Rather than focusing solely on the givenness of information—whether it is given or new—Ariel's model describes the accessibility of referents through flexible and graded terminology. In Ariel's hierarchy (also called the accessibility marking scale), again developed for spoken English, speakers choose referring expressions (REs) not just based on whether the referent is familiar or presupposed but on how accessible that referent is in the discourse. For instance, full-noun phrases are used for low-accessibility referents (first mentions). In contrast, pronouns and zero anaphora are used for high-accessibility referents (which are assumed to be active in the listener's mind). Ariel's hierarchy/accessibility marking scale is given in (2).

(2)

Low Accessibility

Full name + modifier
Full name
Long definite description
Short definite description
Last name
First name
Distal demonstrative + modifier
Proximate demonstrative + modifier
Distal demonstrative + NP
Proximate demonstrative + NP
Stressed pronoun + gesture
Stressed pronoun
Unstressed pronoun
Extremely high accessibility markers (e.g., gaps, reflexives, agreement)

High Accessibility

(Ariel, 1990, p. 73)

In this model, the accessibility status of an antecedent RE is influenced by different factors, such as its distance from the last mention, competition with other referents, its salience, and the unity of the referent, as presented in (3).

(3) Distance: The distance between the antecedent and the anaphor (relevant to subsequent mentions only).
Competition: The number of competitors on the role of antecedent.
Saliency: The antecedent being a salient referent, mainly whether it is a topic or a non-topic.
Unity: The antecedent being within vs. without the same frame/world/point of view/segment or paragraph as the anaphor.

(Ariel, 1990, pp. 28-29)

In summary, according to Ariel, two main factors determine the accessibility of an antecedent in relation to an anaphor. First, some antecedents are naturally more salient and likely to be highly activated in memory, making them more accessible. Second, the type of relationship between the antecedent and an anaphor is important. Physical proximity and coherent connection to the anaphor increase the likelihood that the antecedent remains activated when the anaphor is encountered.

There are also other accounts that link the RE choice and discourse cohesion, like the Centering Theory (Grosz, Joshi, & Weinstein, 1995), which also has computational applications to tasks like anaphora resolution, pronoun resolution, and discourse coherence in dialogue systems, unlike the previously mentioned theories. Grosz mostly focuses on the role of attentional states in discourse and how the selection of REs helps maintain or shift focus between referents, which are called centers. The theory distinguishes between forward-looking centers (topics likely to be referenced in future discourse) and backward-looking centers (which maintain coherence by linking to previous discourse (Walker, Joshi, & Prince, 1997). The latter is akin to the definition of the concept of Topic discussed by Givón (1983), Reinhart (1981), and Horn (1986). According to this understanding, pronouns and other attenuated forms are often used when the referent remains at the center of attention, while more explicit forms signal a shift in focus.

Speakers often use different strategies to distinguish between different accessibility levels. They most often use what is referred to as 'tripartite categorization' to classify discourse contexts of referents and talk about accessibility (Gullberg, 2006; Hickmann, Hendriks, Roland, & Liang, 1996). Referents can be grouped according to their discourse statuses, which are introduced, maintained, and re-introduced. In Gullberg's (2006) annotation, introduction refers to the first mention of a new referent independent of any grammatical role (e.g., for both subjects and objects), whereas maintenance indicates continuing the same referent across one or more clauses. Finally, re-introduction occurs when a previously mentioned referent (i.e., there is at least one intervening clause, including a different referent) needs to be brought back into the discourse. In this classification, introduced referents are expected to be marked with more explicit forms (e.g., nominals), maintained referents with more implicit or reduced forms (e.g., pronouns or zero anaphora), and re-introduced references can be marked with both types of forms[3].

Instead of the aforementioned tripartite classification of discourse contexts, Toole (1996) devised a graded and rule-based scale for English as in (4, 5, 6) to calculate the numerical accessibility level of an entity with scores ranging from -2 (least accessible) to 6 (most accessible) based on the four factors outlined in Ariel's (1990) framework. Toole calculations in these scales were based on theoretical and experimental work in reference tracking (Ariel, 1990; Givon, 1983; Marslen-Wilson, Levy, & Tyler, 1982; Tomlin, 1987). The four factors to be considered are distance and unity, competition, and topicality. Briefly, distance and unity measure how recently a referent was mentioned and how closely it is linked to the current proposition, assigning higher scores for closer mentions (e.g., same or adjacent propositions). Competition accounts for the presence of other entities in the discourse, reducing accessibility when competing referents are introduced between mentions. Finally, topicality evaluates the frequency of a referent's mention within recent propositions, increasing accessibility for more frequent mentions.

[3] Although a referent that has been re-introduced into a discourse is technically already known to the addressee and could be referenced with a pronoun or a zero noun phrase like in the case of maintenance, the cognitive reality is more complex. Chafe (1976) argues that as new information and referents are introduced, the mental accessibility of previously mentioned referents declines by time. This means that the addressee's attention and memory have limits, and older referents can fade from consciousness as newer ones take precedence. So, while some re-introductions may require more explicit forms for reference, some are satisfied with less explicit forms.

(4) Toole's scale of distance and unity
For entity A at point X in the discourse (A_x);
a. If the last mention of A is in the same proposition as A_x then accessibility level of A_x is four.
b. If the last mention of A is in the proposition immediately previous to the proposition of A_x then accessibility level of A_x is three.
c. If the last mention of A is in this episode but not in proposition of A_x or in the proposition previous to A_x then accessibility level of A_x is two.
d. If the last mention of A is in the previous episode to A_x then accessibility level of A_x is one.
e. Otherwise, the accessibility level of A_x is zero.

(Toole, 1996, p. 273)

(5) Toole's scale of competition For entity A at point X in the discourse (A_x);
a. If there are no matching entities between A_x and last mention of A, there is no change to accessibility rating of A_x.
b. If one matching entity has been mentioned between A_x and last mention of A then accessibility of A_x is reduced by one.
c. If more than one matching entity has been mentioned between Ax and last mention of A then accessibility of A_x is reduced by two.

(Toole, 1996, p. 274)

(6) Toole's scale of topicality
For entity A at point X in the discourse (A_x);
a. If A has not been mentioned in the last four propositions[4] then there is no change to accessibility level of A_x.
b. If A has been mentioned once or twice in the last four propositions then accessibility level of A_x is increased by one.
c. If A has been mentioned more than twice in the last four propositions the accessibility level of A_x is increased by two.

(Toole, 1996, p. 274)

Toole applied this methodology to data across four different genres in English, namely, novels, book reviews, casual conversations, and interviews. In her analyses, she grouped accessibility scores from -2 to 0 as low accessibility, scores from 1 to 3 as mid accessibility, and scores from 4 to 6 as high accessibility. Toole also found a strong relationship between accessibility and RE (i.e., more nominal use for lowly accessible referents and more zero anaphora use for highly accessible referents) in the eight texts across the four genres that were analyzed.

[4]Toole reported that the choice of four propositions was arbitrary but it was based on the studies in Givón 1983, which suggested that as the number of intervening propositions increases, the likelihood of needing a more explicit RE (e.g, a nominal) also increases.

As stated before, all the frameworks reviewed above were originally developed for spoken languages. These models, while diverse in terms of their functions, all shared a common focus on the cognitive and pragmatic processes underlying our linguistic choices in spoken communication. The next subsection will shift focus to sign languages, with a particular emphasis on TİD, examining how these theories have been adapted or could be extended to address the unique features of visual-gestural communication systems.

2.2 Establishing discourse cohesion in sign languages

Each language has its own way of tracking referents in discourse. Reference tracking refers to the process by which speakers or signers maintain cohesion in discourse by tracking the entities (referents) being discussed. As explained in the previous subsection, the choice of form is guided by certain cognitive and discursive factors. We know that, like speakers (Ahrenholz, 2005; Aksu-Koç & Nicolopoulou, 2015; Azar & Özyürek, 2015; Debreslioska, Özyürek, Gullberg, & Perniss, 2013; Hickmann & Hendriks, 1999), signers too produce extended discourses and choose from different linguistic forms to track referents by paying attention to their accessibility (Ahlgren & Bergman, 1994; Czubek, 2017; Ferrara et al., 2022; Garcia & Sallandre, 2014; Kibrik & Prozorova, 2007; Morgan, 2000; Perniss & Özyürek, 2015; Schlenker, 2017; Swabey, 2002).

This subsection first describes the REs available for Turkish Sign Language (Türk İşaret Dili - TİD) as well as other sign languages and explains their position on the accessibility scale.

2.2.1 REs in TİD

TİD is the official language of the Deaf community in Türkiye (2005, Disability Law, under the Disability Act subsection no. 5378), and it is believed to have been in use since 1889, making it one of the oldest reported sign languages (Taşçı & Göksel, 2018; Zeshan, 2003). According to the estimated numbers, there are 187.500 to

337.500 deaf people in the country, but it is difficult to estimate how many of them actively use TİD, although the number of TİD signers is believed to correspond to 0.13% to 0.27% of the total population (Kemaloğlu & Kemaloğlu, 2012). The first linguistic work on TİD is considered to be Zeshan's (2003) analysis that examines the phonological, morphosyntactic, and lexical properties of the language. Following Zeshan, more research started to investigate the linguistic structure in TİD (e.g., see Kubuş's 2008 thesis for the phonological properties of TİD; see Açan, 2007; Gökgöz, 2009, 2013, 2024; Makaroğlu, 2013; Göksel, Kelepir, Pfau, Steinbach, and Herrmann, 2016 for work on its morphosyntax; and see Kelepir, 2018; Nuhbalaoğlu, 2018; Keleş et al., 2023 among many others for work on its semantics and pragmatics). More recently, there has been experimental work on the acquisition of TİD, which examine age of acquisition effects on certain components of language, such as morphosyntax, verbal fluency and spatial language acquisition (Karadöller Astarlaroğlu, Sümer, & Özyürek, 2017; Kayabaşı & Gökgöz, 2023; Keleş, Atmaca, & Gökgöz, 2022; Sevgi & Gökgöz, 2023).

The REs that have been attested in sign languages include various types of nominal constructions such as bare nouns (e.g., FARE[5] 'mouse'), fingerspelled nouns (e.g., F-A-R-E), which refers to the signing of the individual letters in the word, and modified nouns (e.g., CAT FEMALE WHITE 'the white female cat'), explicit pronouns (the index sign), as well as some constructions mostly peculiar to sign languages, such as classifier predicates, size and shape specifiers, zero anaphora tools, such as verbal predicates and constructed action. Each will be exemplified in the below subsection, and the examples for each of these categories come from the data collected as part of the research project, Boğaziçi University, Bilimsel Araştırma Projeleri (BAP), entitled "Supporting Sign Language Development of Deaf Children with Hearing Parents through Linguistically Informed Preschool Stories" (Start-up Project, #14458, PI: Kadir Gökgöz), which also constitute the data analyzed and

[5] By convention, sign glosses are given in small caps and translations between single quotation marks.

reported in this thesis. This subsection describes these constructions under five main headings: nominals, pronominals, classifiers, CA, and verbs.

2.2.1.1 Nominals

Nominals are used in TİD extensively to track referents that are generally not easily accessible. They mainly include bare noun phrases, as exemplified in (7), fingerspelled nouns, as in (8), as well as modified noun phrases with adjectives, as in (9), and the index sign, as in (10). Bare noun phrases, such as FARE 'Mouse', can be utilized in TİD to track referents. Additionally, nominals can be fingerspelled, which, as mentioned above, refers to the signing of the individual letters of an alphabet for the surrounding spoken language for a proper noun, foreign or borrowed word, or a concept that does not have an equivalent lexical sign yet. Moreover, when a referent is initially mentioned, it may be followed or preceded by adjectives or numbers that modify the noun (Özsoy & Nuhbalaoğlu, 2014). Definiteness can be indicated by a prenominal or postnominal index sign (KEDI IX), functioning like a demonstrative ('that cat') or a determiner ('the cat'), but the distinction between these two functions is still up for debate (Nuhbalaoglu, 2018).

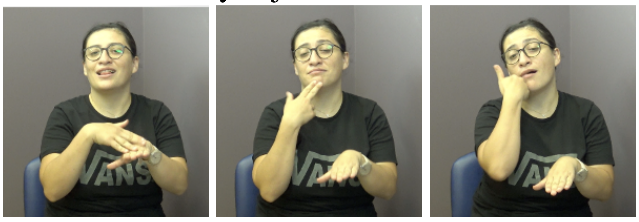
(7) Bare nouns in TİD

MOUSE CAT

(8) Fingerspelled noun in TİD

T -O -M

(9) Noun modified by adjectives in TİD



CAT FEMALE WHITE

’(The) white female cat.’

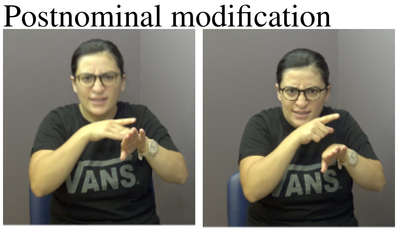
(10) Noun modified by adjectives in TİD

a. Prenominal modification

IX CAT

‘The/this cat’

b. Postnominal modification



CAT IX

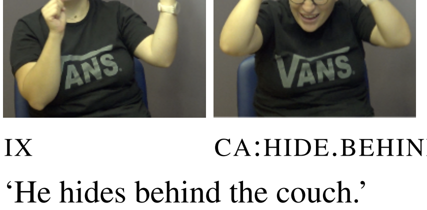
‘The/this cat’ / ‘Cat there’.

2.2.1.2 Pronominal index sign

The index (IX) sign, which is shown via an extended index finger, has been argued to serve various roles in sign languages, including deictic marker, locative indicator, pronominal use (both person and demonstrative), and definitive marker (Bahan, Kegl, MacLaughlin, & Neidle, 1995; Pfau, 2011). Once the referential locus is established for a discourse referent, the pronominal IX, or the extended index finger, might refer to the referent associated with the relevant area, as illustrated in (11). A referential locus is assigned to abstract areas in the signing space by associating a referent to that area through a means such as signing the referent in that area, by looking at that area, shifting the body posture, pointing, or a combination of two or more of these. In (11), the signer is pointing to an abstract place in the signing space, which is associated with the cat.

IX is mainly used in signed languages when re-introducing a topic into discourse (Emmorey, Norman, & O'grady, 1991). This use of IX is also observed in TİD, where it serves the same pronominal function for topic re-introduction (Sevinç, 2006) and is rarely employed for introduction (Perniss & Özyürek, 2015). Studies have also shown that in contexts where referents are highly accessible, such as during the maintenance function of discourse, there are comparatively fewer instances of pronominal IX being used (Cormier, Smith, & Zwets, 2013; Frederiksen & Mayberry, 2016, 2019). As mentioned above, this pronominal use of IX, however, is licensed only if the non-present antecedent is already associated with a spatial locus or if the referent is physically present in the signing area. When the referent is physically present, IX operates deictically by pointing directly to the referent's location in space (Schlenker, 2017). Once the spatial relationship between the referent and the signing space is established, it becomes feasible to re-introduce those referents into the discourse by pointing to their assigned location. Also, note that more recent research has identified what is known as referentially underspecified pronominal IX. This form of IX does not require the non-present antecedent to be anchored to a previously established spatial locus (Nuhbalaoglu, 2018). This contrasts with the traditional understanding of pronominal IX, which hinges on a spatial anchor. In this thesis, I refer to pointing signs that are non-deictic and occur without a noun phrase as pronominal IX, even though some researchers have interpreted such instances as demonstratives (Koulidobrova & Lillo-Martin, 2016).

(11) Pronominal IX in TİD



IX CA:HIDE.BEHIND.COUCH

'He hides behind the couch.'

2.2.1.3 Classifier predicates in TİD

Classifier predicates are complex morphological constructs integral to the grammar of sign languages (Cogill-Koez, 2000; Gökgöz, 2024; Schembri, 2003; Supalla, 1986; Zwitserlood, 2012). These manual and iconic representations depict both animate and inanimate entities, capturing various properties of the referent, such as its location within the signing space, movement patterns, trajectories, size, and shape (Emmorey, 2003; Talmy, 2003). classifier predicates are frequently used in TİD in a similar manner to other sign languages (Dikyuva, Makaroğlu, & Arık, 2015).

Engberg-Pedersen (1993) categorizes classifier predicates into four main types: whole entity classifiers (WECL) represent entire referents, providing a holistic depiction of the entity like humans; extension classifiers, also known as size and shape specifiers (SASSes) (Cormier, Schembri, Vinson, & Orfanidou, 2012; Zwitserlood, 2012), body part classifiers (BPCL), and handling classifiers (HCL). WECLs, as exemplified in (12), represent entire referents, providing a holistic depiction of the entity. SASSes can be classified into tracing and static. Static SASSes (13a) "consist of a handshape (or combination of two hands) that indicates the size/shape of an entity," and tracing SASSes (13b) "have a movement of the hand(s) that outlines an entity's size/shape in which the shape of the manual articulator denotes the dimensionality of that entity" (Zwitserlood, 2012, p. 160). Static specifiers have been argued to be part of classifier predicates, whereas tracing specifiers have been excluded from such predicates (see Kimmelman, Pfau, & Aboh, 2020). In TİD, signers often prefer the I-handshape configuration for tracing two-dimensional geometrical shapes and utilize flat or claw handshapes for three-dimensional shapes like surfaces (Kubuş, 2008).

(12) Whole entity classifiers (WECLs) in TİD

WECL:SITUATED

'(The two mouses) are situated here.'

(13) SASSes in TİD

a. Static

STATIC:HAT EXIST

'There is a hat (shaped like this).'

b. Trace

TRACE:TUNNEL

'Tunnel (shaped vertically like this)'.

BPCLs (14) focus on specific parts of animate beings, such as limbs. HCLs (15) are associated with transitive verbs and illustrate how an object is manipulated or handled by an agent, effectively conveying the manner of the action performed. Within the discourse cohesion framework, HCLs are frequently considered an extension of constructed action (CA) due to their depiction of handling actions performed by animate agents (Cormier, Fenlon, & Schembri, 2015; Dudis, 2004). I also take such constructions as part of CA in this thesis.

(14) Body part classifiers (BPCLs) in TİD

MOUSE ROBOT MOUSE BPCL:STAND

’The mouse and the robot mouse are standing next to each other.’

(15) Handling classifier (HCL) in TİD

HCL:HOLD.MOUSE

‘(The cat) is holding the mouse.’

When we compare pronominal IX to classifiers, which are both implicit in terms of form, some studies also indicate that pronominal IX signs may be dispreferred in favor of classifier constructions (Cormier et al., 2013; Frederiksen & Mayberry, 2019). Conversely, other research reports an abundance of pronominal IX within signers’ discourse (Bel, Ortells, & Morgan, 2015; Morgan, 2000). Ariel’s (1990) scale for accessibility primarily proposes that more attenuated forms are used for highly accessible referents. Therefore, pronominal IX signs are expected to occupy a higher position on the hierarchy than classifiers, given their less explicit nature. Despite this, signers commonly employ classifiers to maintain previous REs. Some studies have also reported sporadic use of classifiers for referent introduction and re-introduction (e.g., Morgan, 2005).

2.2.1.4 Constructed action

Many articulators, including the signer’s head position, facial expressions, torso and arms, and sometimes even their hands, are used to represent the movements and expressions of the referenced character(s) (Smith & Cormier, 2014). Accordingly, the signer adopts the perspective of an entity in the retold story, physically enacting their movements and actions (Metzger, 1995). Signers also may omit the explicit referent anaphor (e.g., omit the nominal form), and they do not assign locations to previously established referents but instead designate their own body as the specified locus when using CA. This change in the spatial scale is often marked with non-manual markers, with a break in eye gaze and facial expression being the most informative cues (Engberg-Pedersen, 1993); however, it is important to note that Janzen (2004) indicates that continued gaze with the addressee can co-occur with CA. In example (16), the construction demonstrates how signers utilize the signing space to imitate the actions of an animate character. In this case, the signer took on the role of the mouse and used their upper right arm and hand to imitate sleeping. CA, therefore, creates the impression that the characters are interacting directly with other characters and objects (e.g., a pillow). The intensity of this construction may also change, resulting in classifications such as subtle CA, reduced CA, and overt CA, the last of which is the strongest form (Jantunen, De Weerdt, Burger, & Puupponen, 2020). This is based on how many limbs have been actively present in the enactment and the degree to which the signer adopts the character's perspective (e.g., nonmanuals signaling the facial expression of the reported character). By this explanation, (16) constitutes an example of an overt CA where the involvement of the non-manuals that contribute to the facial expression of the referent is high. Generally, CA is a tool predominantly used for stories, whereas the percentage of CA is quite low in daily conversations: 39% versus 5% in French Sign Language (Puupponen, Kanto, Wainio, & Jantunen, 2022).

(16) Constructed action in TİD

CA:SLEEP SEE HAT

'(The mouse) was sleeping (like this), and then he saw the hat.'

2.2.1.5 Verbal constructions

Many sign languages permit the omission of subjects and objects, frequently utilizing zero anaphora to refer to entities that are highly accessible. In sign languages, this is typically achieved through the use of space within the signing area, allowing signers to narrate events without the need for overt arguments. In her seminal work on American Sign Language (ASL), Padden (1986) proposed a classification of verbs into groups to better capture how different types of predicates function in discourse. She identified three categories: (i) plain verbs, as in (17), which are uninflected for person or number and thus show no agreement with subjects or objects; (ii) agreement verbs, as in (18), where the directionality of the sign can be adjusted to indicate agreement with specific persons or numbers, using starting and ending points that correspond to various loci in the space; (iii) spatial verbs (e.g., IX$_1$ $_{3a}$PLACE$_{3b}$ CUP 'I placed the cup from this location to that'), which include locational affixes to indicate where an action occurs but, like plain verbs, are not inflected for person or number.

(17) Plain verb in TİD

SICK BECOME

'(The cat) became sick.'

(18) Agreement verb in TİD

Right Hand Gloss: $_{1A}$LOOK$_{3B}$
Left Hand Gloss: HCL:HOLD.MOUSE

'(The cat) looked (at the mouse) while holding (him).'

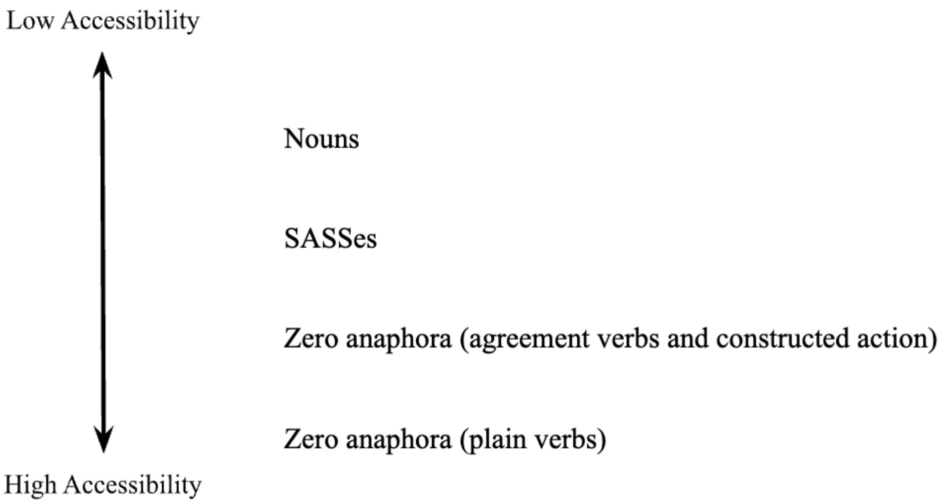
In TİD, null arguments are too widely permitted. Following Padden, Sevinç (2006) noted that while argument omission is common in TİD, there are certain constraints regarding null marking for agreeing and plain verbs. However, recent studies challenge Sevinç's earlier findings. Kayabaşı et al. (2020), for instance, suggest that the omission of overt arguments in TİD is more flexible than previously understood, as signers appear to drop arguments across all verb types without the strict limitations that were initially proposed.

2.2.2 Reference tracking studies in the visual-manual modality

The tenants of both formal and cognitive-functionalist frameworks in the sign language reference literature have ubiquitously found that it is possible to explain the patterns for building a cohesive discourse (e.g., the choice of REs) with the idea of cognitive accessibility (Ariel, 1990; Gundel et al., 1993). As such, across many different sign languages, signers track new referents with phonetically heavier and more informative tools such as bare nouns and maintain a previous referent with tools of zero anaphora, such as classifier predicates, constructed action and verbs (Ahlgren & Bergman, 1994; Barberà, 2013; Bel et al., 2015; Cormier et al., 2013; Ferrara et al., 2022; Garcia & Sallandre, 2014; Hodge, Ferrara, & Anible, 2019; McKee, Schembri, McKee, & Johnston, 2011; Morgan, 2000; Perniss & Özyürek, 2015; Pizzuto, Rossini, Sallandre, & Wilkinson, 2006; Swabey, 2002; Wulf, Dudis, Bayley, & Lucas, 2002). To explain the distribution of available referring forms, most of these studies alluded to Ariel's (1990) Accessibility Hierarchy, with the exception of Swabey (2002), who based it on Gundel's (1993) Givenness Hierarchy, and Barberà & Massó (2009) who used Grosz's (1995) Centering Theory. Moreover, most of these analyses grouped REs under general categories (e.g., nominals, pronominals, classifiers, and null constructions).

Swabey (2002) investigated how referents are tracked in American Sign Language (ASL) and found both similarities and differences with English. Her study revealed that ASL signers commonly use bare nouns to introduce or re-introduce referents in discourse, while zero anaphora or verbal constructions are favored for referents that are already well-established or accessible to the addressee. Overt referential expressions, such as pronouns or full noun phrases, tend to be avoided when the referent is "in focus," reflecting adherence to the Gricean maxim of quantity, where more concise forms are preferred. ASL and English pattern very similar with respect to the status of activation required for zero anaphora and pronominal elements, but there were also main differences. While ASL signers preferred to use zero anaphora (e.g., classifiers, verbs, and constructed action) for reference maintenance, English speakers almost always relied on pronouns and nouns to make a reference to “in-focus” entities. There are also certain modality-specific differences. For example, ASL signers can capitalize on the signing space to refer to multiple entities simultaneously. However, Swabey’s general conclusion was that native signers use language economically with the least amount of effort, as previous cognitive-functionalist theories claimed. Also, Frederiksen and Mayberry (2016, 2019) conducted a story-telling experiment on ASL to investigate the forms by which signers tracked referents and proposed the hierarchy in (19) for ASL:

(19)



(Frederiksen & Mayberry, 2016, p. 17)

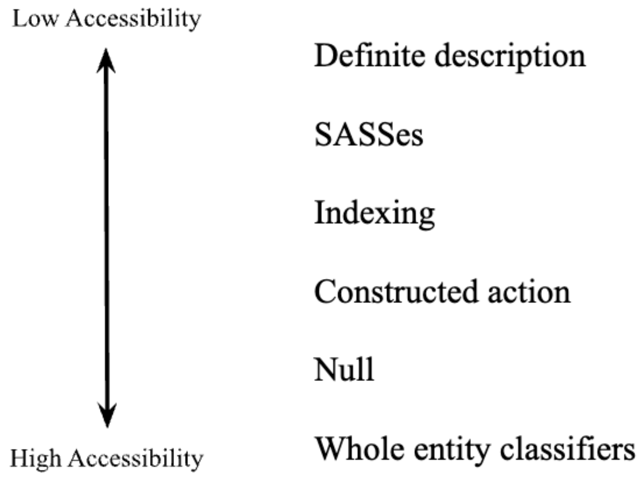
The authors reported similar patterns of form in ASL observed in spoken and other sign languages. Signers primarily used nominals (and secondarily SASSes) for introduction and zero anaphora for maintenance (69% across all discourse contexts) as well as re-introduction. To continue the previous referent, ASL signers used zero anaphora (agreement verb, CA, plain verb) the most, followed by whole entity classifiers. The most preferred zero anaphora types were CA and agreement verbs. Only 7% of observed REs belonged to a re-introduction, and signers used mostly zero anaphora (67%) and then nominal (33%) for bringing a referent back to discourse. Classifiers were mainly for reference maintenance and introduction but not for re-introduction. Surprisingly, the authors failed to find any pronominal use (e.g., a pointing sign that functions as a person pronoun) in their study. Few other studies also reported very sporadic uses for pronouns in signed narratives (Czubek, 2017; Swabey, 2002 for ASL; Hodge et al., 2019 for Australian Sign Language; Keleş et al., 2023 for TİD) but this finding was not compatible with other sign languages, such as Catalan Sign Language (Llengua de Signes Catalana - LSC) (Bel et al., 2015), British Sign Language (BSL) (Morgan, 2005), and German Sign Language (Deutsche Gebardensprache - DGS) (Perniss & Özyürek, 2015).

Czubek (2017) expanded Frederiksen and Mayberry's (2016) accessibility hierarchy by investigating simple and complex ASL narrative productions of 19 deaf adult native signers. A story-telling experiment was conducted with ASL folklore narratives and the Balloon Story (a well-known 6-panel wordless elicitation story used in the literature) (Karmiloff-Smith, 1979). Czubek used Toole's (1996) scale (based on Ariel's 1990 principles) and quantified accessibility scores for each RE type identified in ASL. He also distinguished between REs used in isolation (20) and REs realized as constellations (21), that is, REs that co-occur, and proposed separate hierarchies for each case.

While the hierarchy for REs in isolation successfully replicated the one proposed by Frederiksen and Mayberry (2016), the hierarchy for REs in constellation also contributes another layer to possible reference tracking tools in ASL. For referents found in isolation and in the Balloon Narratives, Czubek observed that the percentage of re-introductions in his study was higher (19%) compared to

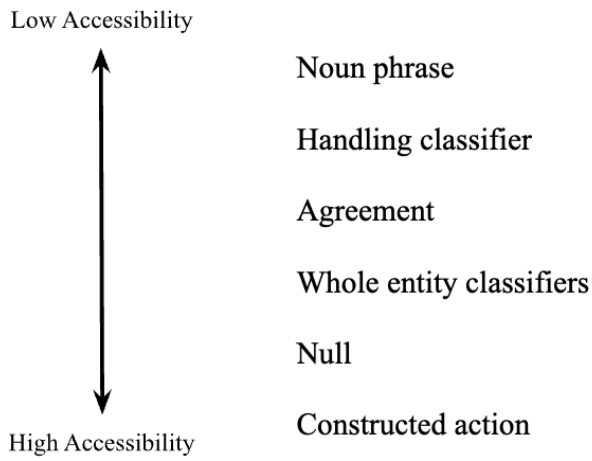
Frederiksen and Mayberry (7%). Introduced referents constituted 17% of all REs, while the maintained REs had the highest percentage (64%).

(20) Czubek's proposed accessibility hierarchy in ASL for REs in isolation.



(Czubek, 2017, p. 151)

(21) Czubek's proposed accessibility hierarchy in ASL for REs in the constellation.



(Czubek, 2017, p. 151)

Unlike Frederiksen and Mayberry, Czubek categorized CA and verbs separately. Noun phrases were the most common type for introduction (94%) and re-introduction (63%). CA was less likely to occur as an introduction (1%) and re-introduction (19%) compared to maintenance contexts (39% compared to 16% in Frederiksen and Mayberry). The pronominal index occurred very infrequently, similar to Frederiksen and Mayberry, and most of the index signs accompanied nouns. Czubek also attested handling classifiers in the data, and their frequency was a function of the narrative stimuli, correlating with object entities. As for the accessibility scores, Czubek showed that noun phrases (followed by SASSes and pronouns) in ASL have the lowest accessibility values among all RE types, indicative of their highly explicit nature. Whole entity classifiers had the highest accessibility values, and they were followed by CA, agreement verbs, and null markers. The ASL literature stimuli (which is more complex than the first narrative) also had similar effects, with the exception that whole entity classifiers and CA had statistically lower accessibility compared to the Balloon Story. While the whole entity classifiers had the highest accessibility in Czubek's study, the behavior of such constructions seems to be language-specific given Morgan's (2005) and Keleş et al.'s (2023) observations for BSL and TİD, respectively. In these studies, such classifiers had moderate explicitness or accessibility status, given their position between nominals and CA.

Other studies examining a variety of sign languages like Australian Sign Language (Auslan) (Hodge et al., 2019; McKee et al., 2011), BSL (Morgan, 2005), DGS (Perniss & Özyürek, 2015), French Sign Language (Langue des Signes Française - LSF) (Garcia & Sallandre, 2014), LSC (Bel et al., 2015) and TİD (Keleş, Atmaca, & Gökgöz, 2023) have reported similar patterns in referential strategies.

For example, Perniss and Özyürek (2015) investigated the visual-manual modality in terms of reference tracking by comparing overt REs found in DGS, German, and German co-speech gestures. Like Swabey, Frederiksen and Mayberry, and Czubek, the authors also found support for the predictions of referential accessibility principles (i.e., the quantity of marking on form changes with accessibility), independent of language modality. However, German used more overt forms for reference maintenance and re-introduction compared to DGS, which had a higher occurrence of null subjects. The authors also find that co-speech gestures and signs use spatial modification as a modality-specific feature, creating a distinction between unimodal and bimodal systems in terms of visual cohesion.

Hodge et al. (2019) used clustering and mixed effects modeling to examine discourse cohesion in Auslan, with the finding that accessibility or information density is a significant predictor for RE choice (e.g., fuller forms for less accessible referents). In addition, the authors looked at the role of animacy (animate versus inanimate) and humanness (human versus non-human) and found that the RE choice significantly depended on these two factors. As an example, human referents received fewer RE forms than animals and then inanimate referents.

Ferrara et al. (2022) analyzed a multilingual narrative corpus that consisted of videos from Auslan, Irish Sign Language, Finnish Sign Language, Norwegian Sign Language, and Swedish Sign Language. Their findings supported previous conclusions, showing that bare nouns are predominantly used to introduce new referents, while reduced forms like verbs are employed for maintaining reference once established. Moreover, all these languages made a bipartite distinction between discourse context (introduced versus non-introduced) and animacy (animate versus inanimate). The cross-linguistic differences, however, were mostly observed in the RE choices across accessibility and animacy categories and fingerspelled words.

Finally, in an earlier study, we examined the strategies of tracking entities in narratives among deaf adult signers in TİD (Keleş et al., 2023) and proposed the following hierarchy in (22). Unlike Czubek, we did not use quantified accessibility scores in this study, but we had a more extensive classification of REs. TİD signers also produced a very low number of pronominal IX in this study (40 in total out of 4007 observations), aligned with ASL and Auslan. Furthermore, there were 1031 occurrences (26%) of referent introduction, 1589 occurrences (40%) of referent maintenance, and 1387 occurrences (34%) of referent re-introduction in the data. Among the classifiers, whole entity classifiers were the most frequently used type, followed by body part classifiers and then SASSes.

(22) Keleş et al.'s proposed accessibility hierarchy in TİD.

Low Accessibility

Modified nominal (with adjectives or pointing)

Bare nominal

SASSes

Whole entity and body part classifiers

Constructed action

Agreement verb

Plain verb

High Accessibility

(Keleş et al., 2023, p. 27).

For first mentions, signers predominantly chose SASS, while whole entity classifiers were more commonly employed for referent maintenance and re-introduction compared to body part classifiers and SASS. CA was the most frequently used zero anaphora tool, and it was then followed by verbal constructions like plain verbs and agreement verbs. TİD signers used zero anaphora very sporadically markers when introducing a referent, with CA being the most common. Additionally, CA was employed more frequently for re-introduction than for maintenance, unlike Frederiksen and Mayberry, and Czubek. The use of agreement verbs was similar for both maintained and re-introduced referents, whereas signers preferred to use plain verbs more often in maintenance contexts than in re-introductions.

2.3 Age of acquisition effects

The amount of work dedicated to the acquisition track of certain structures of sign languages among deaf signers has flourished in recent years, and there are many reasons for this growing interest. Of course, the native acquisition of sign languages (i.e., deaf or hearing children learning how to sign from a deaf caregiver) parallels that of spoken languages, and they both display similar developmental stages (Meier & Newport, 1990; Woll, Meurant, Sinte, Van Herreweghe, & Vermeerbergen, 2013). However, this is not always the case; in fact, only a minority (10%) of deaf children are born to deaf parents, and consequently, the majority of signers (90%) are born into hearing parents who have little or no experience with a full-fledged sign language (Mitchell & Karchmer, 2004). This engenders a modality mismatch in the interaction of children and caregivers, resulting in a delay in first language acquisition. In such acquisition environments, the first frequent and systematic exposure to signs is often delayed until these children receive formal education (Mayberry, 2007). Many studies have reported that the dearth of conventional linguistic input and late exposure to signs in the early few years negatively affect different components of language production and processing even when these children reach adulthood. Among the reported undesirable consequences for the deaf adult late signers (deaf children of hearing parents), complex morphosyntactic structures are sensitive to the effects of first language delay (Boudreault & Mayberry, 2006; Cheng & Mayberry, 2021; Kayabaşı & Gökgöz, 2023; Sevgi & Gökgöz, 2023).

As for discourse cohesion, there are only a few investigations that report on age of acquisition effects for delayed L1 acquisition among deaf adults (Keleş et al., 2023) as well as adult L2 acquisition (Bel et al., 2015; Frederiksen & Mayberry, 2019), suggesting over-explicitness akin to the findings for second language (L2) speakers. Reference tracking or building a cohesive discourse is challenging for second-language learners since it demands the simultaneous integration of syntactic and discursive properties. One explanation that is put forward about the complexity of syntax and discourse interaction is the Interface Hypothesis (Sorace, 2011; Sorace & Filiaci, 2006). Numerous other studies have tackled the issue of attaining narrative competence in acquiring a second or foreign language. Overall, the findings indicate that at earlier stages of language learning, children tend to show under-explicitness (e.g., using less form even for inaccessible referents) but intermediate and advanced L2 learners display over-explicitness instead (e.g., using more form than necessary for accessible referents) (Ahrenholz, 2005; Hickmann & Hendriks, 1999; Leclercq & Lenart, 2013; Williams, 1988). Nonetheless, over-explicitness has been argued to vary heavily among learners, and it can change due to the learners' different linguistic backgrounds. Some researchers attribute this behavior among L2 learners to having more clarification statements and including redundant references (J. Ryan, 2015). Another attempt at explaining this comes from Hendricks (2003), who claimed that over-explicitness might stem from complex noun phrase types and be tied to a lack of sufficient pragmatics and discourse-level skills that enable learners to leverage between accessibility and phonetic form.

Bel et al. (2015) compared the reference productions of deaf adult signers (*N* = 11) to hearing L2 learners (*N* = 13) in LSC narratives. Their results showed that hearing signers tended to use more explicit REs, such as full nouns or pronouns. Comparably, deaf signers produced more implicit forms, or in other words, they made references with less phonetic form. However, the results did not show a statistically significant interaction between discourse status (i.e., whether the referent had been previously mentioned or was new) and the acquisition group (i.e., hearing versus native deaf signers). An important factor to note is the composition of the control group, which included both deaf signers from two groups: those with deaf parents (who acquired sign language from birth) and those with hearing parents (whose exposure to sign language began before age 5). Specifically, out of the 11 native deaf L1 signers in the control group, 5 had deaf parents and natively learned Catalan Sign Language (LSC), whereas the other participants had hearing caregivers and started learning LSC before age 5.

In line with the previous findings, Frederiksen and Mayberry (2019) conducted a study that compared deaf adults, who are native signers of ASL, with hearing second language (L2) learners of ASL. Their results indicated a meaningful difference between these two acquisition groups, specifically in how they used classifiers in re-introduced discourses. Hearing learners were found to use more overt forms, such as nominals, and they produced little to no anaphora. Frederiksen and Mayberry attributed L2 learners' use of more nominals and less zero anaphora to a tendency towards over-explicitness. It is also important to note that the control group that Frederiksen and Mayberry had consisted of signers that had hearing parents, unlike Bel et al. (2015), all of them were still considered to be native signers of ASL because they were exposed to a sign language from birth through their older deaf siblings.

Unfortunately, there is a paucity of research directly comparing deaf signers with native acquisition of a sign language (henceforth, native signers) and those with delayed exposure (henceforth, late signers). To name a few, Becker (2009) examined DGS narratives and deaf children's reference tracking. She found that late signers used more explicit forms like nominals to refer to story characters compared to native signing children, indicating that delayed first language acquisition might lead to more inhibited morphosyntactic skills. Similarly, Cormier et al. (2013) found late signers in BSL, contra natives, were more redundant in maintaining reference, using explicit forms rather than implicit zero anaphora. However, some studies showed that late acquisition does not affect the strategies of referent introduction (Gür, 2024; Gür & Sümer, 2022).

2.4 Ease of articulatory effort and phonetic reduction

Sign language phonetics and phonology are shaped by the anatomy of the hands, arms, and face, which limits and influences how signs are formed and organized. The physical and phonetic aspects of each sign, such as handshape, location, movement, and orientation, are structured based on what the body can comfortably and consistently produce. The structure of signs is shaped by the natural capabilities and limitations of the hands, arms, and other articulators, creating a direct connection between the physical body and the linguistic elements produced (Crasborn, 2012; Stamp, Dachkovsky, Hel-Or, Cohe, & Sandler, 2024). Crasborn cites several studies that have explored how the anatomy of the hands and arms affects the articulation of signs in sign languages. Mandel (1979), for example, identified that the extensor muscles in our fingers do not allow for full flexion of all finger joints when the wrist is also bent as far as it can go because of its length. This anatomical limitation influences the range of motion for both the fingers and the wrist, restricting their simultaneous movement. A practical demonstration of this can be seen when holding the forearm horizontally with the palm facing downward: as a fist is formed rapidly, the wrist extends naturally to support the hand's movement. Similarly, when the wrist bends forward from a neutral position, the fingers extend automatically to accommodate the change in wrist angle, maintaining stability and comfort.

While slower movements allow greater control, the anatomical constraints ultimately limit the extent to which the fingers and wrist can move together. For example, when all fingers close, the wrist tends to extend, resulting in a backward movement of the hand. As such, several other studies reported that distributional patterns of handshapes (e.g., why certain handshapes are more common and easier to do than others), and also other phonological parameters constituting a sign, are tangent to the anatomical and physiological properties of the articulators (Ann, 2006). Napoli et al. (2014) explain the nonuniform distribution of articulatory components of signs (e.g., handshapes) with the notion of articulatory ease or reduction of biomechanical effort, which favors certain configurations over others, making them marked. Sanders and Napoli (2016) make a further distinction between what they term active and reactive effort. Active effort refers to the muscle activation required to produce a specific articulatory configuration, such as forming a particular handshape or orienting the hand in a certain way. Reactive effort, on the other hand, pertains to the physical resistance the signer encounters, like gravitational forces or incidental movements, by studying the lexicons of three different and genetically unrelated sign languages, Sanders and Napoli, in their study, compared stable signs (that do not induce torso movement and thus do not necessitate reactive effort) and destabilizing signs (that induce torso movement and thus require reactive effort). They found that stable signs were favored, and destabilizing signs were disfavored across all three languages. The authors then concluded that ease of articulatory effort does not only contain minimizing active effort but should also be extended to the reduction of reactive effort.

Widely described for spoken languages (e.g., Daniloff & Hammarberg, 1973; Moon & Lindblom, 1994; Recasens, Pallarès, & Fontdevila, 1997), coarticulation involves the blending or overlapping of articulatory movements due to adjacent segments. There is also considerable support for coarticulation effects for phonological parameters, such as location and handshape in sign languages (Grosvald, 2009; Grosvald & Corina, 2012; Mauk, 2003; Ormel, Crasborn, Kootstra, & De Meijer, 2017). Figure 1 shows Grosvald and Corina's (2012) illustration of such coarticulatory effects. Just as various preceding or following vowels can influence the position of the neutral vowel schwa, preceding or following signs at various locations can affect a sign located at the neutral signing space, indicative of coarticulation.

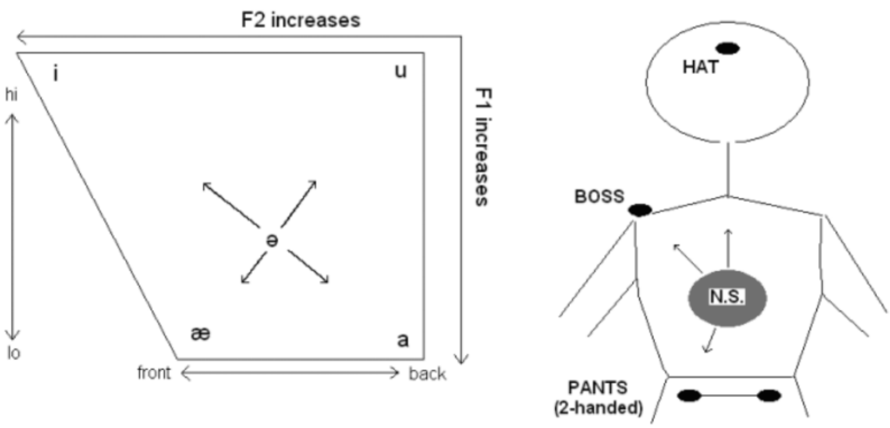


Figure 1. Grosvald and Corina's (2012, p. 39) illustration of coarticulatory effects of different vowels on schwa and different sign locations on neutral signing space

Phonetic reduction, on the other hand, refers to a decrease in the amplitude or size of an articulator, often paired with a shortened duration of the articulator sequence in connected speech or sign. For example, Philip Lieberman (1963) demonstrated that English speakers phonetically reduced a word (e.g., the word nine in his experiment) in the latter occurrences of the word when it became contextually more predictable due to increasing syntactic, semantic, and discursive cues. Reduction, in this case, meant shorter duration and less intelligibility. Phonetic reduction differs from the previously established notion of coarticulation since it generally encompasses not only the shortening of duration but also the articulatory undershoot. Undershoot is a phenomenon that occurs when an articulator does not fully reach its intended target position (Mauk, Lindblom, & Meier, 2008). This results in a less precise articulation of the sound, which often occurs in rapid or casual speech, and this may bring about more coarticulation, consonant lenition, and centralization effects (Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Bybee, 2003; Gahl, Yao, & Johnson, 2012; Jurafsky, 2003). Extending this idea to discourse cohesion, it has been well observed for spoken languages that repeated references (i.e., referents that contain old information or that are maintained or re-introduced) undergo acoustic reduction and shortening (Aylett & Turk, 2004; Bard et al., 2000; Fowler, 1988).

Returning to the visual-manual modality, Tyrone and Mauk (2012) exemplify phonetic reduction in signs with Lucas et al.'s (2002) observation that certain ASL signs around the forehead can shift downward to make contact with the cheek in more fluent signing. For instance, the sign KNOW usually touches the forehead, but often, a phonetic lowering is observed during casual conversations. The authors argued that lowering is less likely to occur if the preceding sign was articulated in a higher location. In an earlier motion capture study by Tyrone and Mauk (2010), ASL signs made in front of the forehead even without physical contact (e.g., WONDER) were observed to reduce phonetically (move to the cheek) due to the rate of signing or the surrounding phonetic context. In these studies, the authors asked the participants to sign WONDER at three different speeds, and lowering was mainly observed in the higher speed conditions. The location of the preceding sign was also found to be an important factor in determining the likelihood of the phonetic reduction. As can be seen from these examples, in sign language production, movements of a sign in fluent signing can be smaller than their citation form. Also, Russell et al. (2011) noticed that frequency plays a role in the lowering of the sign: the more frequently a sign occurs, the more likely it is to be lowered and to greater extents. This also parallels the findings for spoken languages: more frequent words tend to exhibit more phonetic reduction (Bell et al., 2009; Bybee, 2003; Gahl et al., 2012; Lin, Beddor, & Coetzee, 2014). While the causes of this type of frequency-reduction correlation can be attributed to the principle of least effort (Zipf, 1949), usage-based linguistics highlights the routinization of production (Bybee, 2003; Lepic, 2019). According to this view, more practiced words take less time to produce.

Moreover, gestures accompanying frequent words and gestures for an action representing given or old information are known to be simpler in form and less informative (Gerwing & Bavelas, 2005). In contrast, speakers use larger gestures for unpredictable referents (Levy & McNeill, 1992). As such, phonetic reduction is present not only in spoken and sign languages unimodally. It is also found in multimodal communication, as can be evidenced by co-speech gestures.

The tendency to shorten form by a function of predictability or information density extends to signs (Hoetjes, Krahmer, & Swerts, 2014), who investigated the production of referents by signers of Sign Language of the Netherlands (Nederlandse Gebarentaal - NGT) (*N* = 14). Hoetjes et al.'s participants took part in a director-matcher task where the participants depicted figures to an addressee, and the researchers examined their repeated referents, expecting to find a reduction. Hoetjes et al. found evidence for a systematic reduction in repeated referents. Compared to initial mentions, subsequent mentions of referents were considerably shorter.

Going back to the age of acquisition effects, there is only a dearth of observations regarding how native and late signers might phonetically differ in narratives. Börstell et al. (2024) presented data from three sign languages, BSL, NGT, and Swedish Sign Language (Svenskt teckenspråk - STS), across 288 signers and examined whether sociolinguistic factors (age, gender, region, age of acquisition, and family status) as well as token frequency affect the duration of the sign and production rate or not. Their dataset included 50 hours of annotated videos from narratives and conversations. In their study, age and token frequency were found to significantly affect sign duration and rate. Older signers had longer sign duration and lower production rates compared to younger signers. Furthermore, signs with lower frequency had longer duration and lower rate. Other factors like age of acquisition, region (with the exception of BSL), and gender did not successfully predict duration and production rate.

Braem (1999) also analyzed personal narratives produced by early (N = 3) and late (N = 3) signers of Swiss German Sign Language (DSGS) and discovered certain differences between the groups in their rhythmic temporal patterns. All their early signers were prelingually deaf and acquired DSGS from their deaf caregivers. In contrast, the late signers learned DSGS later in life. The authors found that late learners, contra early signers, were usually slower in speed, had a smaller amplitude of movement, and had more holds inside a movement.

2.5 Applying computer vision to sign languages

Thanks to the advancements in computer vision, it is possible to analyze body and facial movements in gesture and sign either with motion capture tools or through pose estimation tracking of live or pre-recorded videos (Börstell, 2023). Kinematic studies (i.e., those pertaining to the investigation of motion and movement) can be classified into two based on whether they use marker-based or markerless solutions (Needham et al., 2021). Marker-based methodologies use censored motion capture, in which the pose of articulators can be tracked with markers or sensors on the body and infrared cameras, and animations are calculated based on the constellation of these markers. Previous research incorporated these marker-based methods to detect sign lowering and phonetic reduction (Tyrone & Mauk, 2010, 2012), create 3D datasets for the annotation of phonetic properties (Jedlička, Krňoul, Kanis, & Železný, 2020; Ranum, Otterspeer, Andersen, Belleman, & Roelofsen, 2024), combined it with eye tracking data to examine CA in signed narratives (Jantunen, Puupponen, & Burger, 2020; Puupponen, Wainio, Burger, & Jantunen, 2015), and investigated phonetic properties like velocity and acceleration (Jantunen, Burger, De Weerdt, Seilola, & Wainio, 2012), as well as telic and atelic verbs (Malaia & Wilbur, 2012). Jantunen et al.

(2020), for instance, analyzed three subtypes of CA, which are overt, reduced, and subtle, in Finnish Sign Language (FinSL) with marker-based motion capture and reported faster and larger movement of the torso and head for overt and reduced CA, and more reduced movement for regular narration and subtle CA.

Instead, markerless pose estimation algorithms use artificial intelligence to track the position and orientation of body parts and yield 2D or 3D pose estimations directly from live or pre-recorded video or image data. Reconstructing 3D motion using indirect methods, such as pose estimation from 2D video or depth cameras, the results are generally less precise compared to the accuracy that can be achieved with direct 3D motion capture techniques, like the aforementioned motion capture technology that use markers on the body (Ranum et al., 2024). That said, markerless pose estimation does not require physical markers to be placed on the body, making it noninvasive, easy to conduct, and cost-efficient (Colyer, Evans, Cosker, & Salo, 2018). Some important open-source advances in the development of markerless pose estimation are OpenPose (Cao, Simon, Wei, & Sheikh, 2017) developed by Carnegie Mellon University, OpenFace (Baltrusaitis, Zadeh, Lim, & Morency, 2018), and more recently, Google's MediaPipe (Lugaresi et al., 2019). In essence, these programs output 3D tracking data from 2D video or image, with one caveat that the estimated depth information (e.g., the z-axis) might not be very reliable (Stamp, Dachkovsky, et al., 2024). Previous research employed such markerless programs to analyze many phonetic aspects of manual articulators across many different sign languages, including sign location analysis, movement patterns, hand dominance, and the size and duration of iconic constructions (Börstell, 2023; Börstell & Lepic, 2020; Fragkiadakis, 2022; Kimmelman & Teresė, 2023; Martínez-Guevara & Curiel, 2024; Normoyle et al., 2022; Östling, Börstell, & Courtaux, 2018; Roh, Lee, Hwang, Cho, & Park, 2024; Slonimska et al., 2024) as well as non-manual features, such as mouthing recognition (Sáenz, 2022), and presence of grammatical eye-brow raising (Kimmelman, Imashev, Mukushev, & Sandygulova, 2020; Kuznetsova, Imashev, Mukushev, Sandygulova, & Kimmelman, 2021) and head nods (Bauer, Kuder, Schulder, & Schepens, 2024). As for TİD, few studies employed open tools like OpenPose and MediaPipe to investigate phonetic correlates of different components of grammar, including telic versus atelic verbs (Yaşar, Kisbet, Şahin, & Gökgöz, 2024), pointing signs (Eroğlu, Şahin, Kisbet, & Gökgöz, 2024), complex motion events and first language delay effects (Gökgöz & Keleş, 2023), and iconicity of a sign (Keleş, 2024). Furthermore, a phonological framework entitled Sign Language Feature Extraction (SL-FE) for TİD was proposed using Google's MediaPipe (Şahin & Gökgöz, 2024).

In addition to these open-source programs, some studies employ Microsoft Kinect, which is a hardware-based system providing 3D body tracking to record signs and gestures in real-time (Zhang, 2012). As such, there have been few studies using this tool for a motion capture analysis in sign language (see Flaherty et al., 2023 for signing space analysis; Krebs et al., (Krebs et al., 2024) for verb and adjective types; Stamp et al., 2024, for classification of different RE types; Stamp, Dachkovsky, et al., 2024, for phonetic reduction in repeated referents).

Several research efforts have applied these markerless pose estimation and motion capture tools to analyze phonetic reduction effects in continuous discourse. Flaherty et al. (2023) used a markerless Microsoft Kinect motion tracker to investigate the signing space size among 17 deaf adult signers of Nicaraguan Sign Language (NSL), which has been argued to emerge in the 1970s when the Nicaraguan government established schools for deaf children. Before this, deaf individuals in Nicaragua lived relatively isolated lives and used home sign, primitive gesture-based communication systems that are not full-fledged languages (Goldin-Meadow, 2005; Özyürek, Furman, & Goldin-Meadow, 2015). When these children were brought together in the schools, they started developing a shared sign language over the last 40 years (Senghas & Coppola, 2001; Senghas, Kita, & Özyürek, 2004). Each new generation of children entering the signing community and acquiring NSL has been documented to add structural complexity to the newly emerging language. While the first cohort mainly depended on gesture-like properties and their individual homesign with a consistent lexicon, the other cohorts added further complexity to NSL (e.g., by introducing spatial morphosyntactic marking and tools for describing spatial relations) (Pyers & Senghas, 2009). It is also well observed that established and urban sign languages like ASL are articulated in a comparatively smaller signing space compared to newly emerging sign languages 2007. Building on this background, Flaherty et al. investigated the signing space of NSL signers with differing years of entry into the signing community with a motion-track camera. The participants described 36 short vignettes (actions between two characters). The authors reported three measures of the space but focused on the "measure of the signing space which takes the Euclidean distance between the position of the wrists in each frame and a fixed point on the body (the Kinect tracked joint between the shoulders)" (Flaherty et al., 2023, p. 11). The results revealed that younger NSL signers significantly used a more restricted signing space than older NSL signers, documenting an ongoing change in signing space size in a newly emerging language.

Testing the predictions of cognitive theories such as principle of least effort (Zipf, 1949) and accessibility hierarchy (Ariel, 1990), Stamp, Dachkovsky, et al. (2024) used Microsoft Kinect motion tracker to examine reduction in the narratives of Israeli Sign Language (ISL), depicted as a young sign language, by comparing the phonetic properties in three nominal REs (CHARLIE-CHAPLIN, LION, and WOMAN) across different discourse contexts (Introduction, Maintenance, Re-introduction). They reported data from 13 native deaf adult signers of ISL (aged between 25 and 76). The authors devised the following six kinematic/phonetic measures to detect reduction effects presented in (23).

(23) Stamp, Dachkovsky, et al.'s (2024) kinematic measures.

1. Duration (D) – the time elapsed in seconds between the first and last frames in the sequence.
2. Distance Covered – the distance in meters, traversed by the skeleton joint p during the sequence. It is computed by accumulating the Euclidean distance between joint location in consecutive frames.

3. Average speed – computed as the Distance Covered by joint p divided by the duration. speed is given in m/s.
4. Variance – the variance of the 3D position of the joint p across the sequence (meters).
5. Volume – the 3D volume of the space (meters) traversed by joint p during the sequence. It is computed by determining the volume of the smallest convex polygon that bounds all 3D locations of the joint p throughout the sequence.
6. Mean distance from body plane – the distance in meters of the skeleton joint from the body plane. The body plane is computed as the 2D plane spanned by three skeleton joints: Shoulder-Left, Shoulder-Right and Spine-Base.

(Stamp et al, 2024, p. 6)

The phonetic measures were collected from three selected joints: (i) the head joint to analyze head movements, (ii) the dominant hand joint (right or left) to analyze hands, and (iii) the spine-shoulder joint to analyze the torso. The authors recruited 15 deaf adult ISL signers (two of them were excluded) in the study and instructed them to watch and then narrate a silent movie to an interlocutor. This resulted in 211 REs (Introduction: 38, Maintenance: 24, Re-introduction: 149). Overall, Stamp, Dachkovsky, et al. found support for phonetic reduction in signed narratives in ISL. They reported a significant reduction in phonetic parameters related to duration, length, and size but not for the ones related to speed and/or distance from the body. However, the detected reduction persisted across different joints or articulators. Overall, nominal REs in ISL were reduced in duration, length, and size when maintained across one or more clauses. In contrast, introduced REs were longer and larger in size. The authors attributed these findings to ease of articulation, which could be explained by Napoli et al.'s (2014) concept of effort reduction.

In an earlier study, Stamp et al. (2024) investigated the kinematic or phonetic properties of different REs (nominals or lexical signs, classifiers, and CA) in ISL with the same motion capture technology. Using the same six measures across three articulators (hand, face, and torso), they found that nominals had a shorter duration, were faster, smaller in terms of distance, volume, and variance, and had a narrower signing space, compared to other REs. Another finding was that CA and classifier had similar kinematic properties despite the fact that both constructions structurally differ. Stamp et al. conclude by highlighting that initial kinematic information can be used to distinguish among these three forms.

More recently, computer vision has been utilized to analyze the phonetic properties of TİD, as well. Yaşar et al. (2024) investigated telicity, whether an action has a clear endpoint, in TİD with pose estimation and found distinct movement patterns (smooth motion for atelic events versus non-harmonic peaks for telic events). Eroğlu et al. (2024) explored the functions of pointing (IX) signs in TİD, focusing on distinguishing between different uses such as pronouns, demonstratives, and locatives. By using MediaPipe's pose estimation, the authors analyzed the phonetic properties of such pointing signs and reported that categories with more features, like locatives, had richer phonological specifications, utilizing spatial parameters more fully compared to others like demonstratives. In a pilot study, we also analyzed the age of acquisition effects in complex motion events in TİD with OpenPose (Gökgöz & Keleş, 2023). We calculated the estimation of articulatory energy by assigning different weights to joints on the body and found that deaf adult late signers in TİD spent more articulatory energy with the non-dominant side of the body compared to native signers. We attributed the findings of this research to Sanders and Napoli's (2016) idea of reactive effort and concluded that native signers of TİD were more successful in inhibiting their non-dominant or left side of the body. However, the caveat associated with this study is that we analyzed these complex motion events in isolation and not inside an extended discourse or dialogue. More relevantly, we also examined the phonetic and iconic characteristics of REs in TİD narratives using OpenPose (Keleş, 2024). In the first experiment, we investigated the amount of movement in different RE forms, finding that forms like Tracing and CA occupied more signing space compared to CL, Verbs, and Nominals, aligning with the findings of Stamp et al. (2024). In the second experiment, we recruited 52 non-signing participants and had them rate the iconicity of 50 REs, revealing that more iconic signs, such as those involving CA or Tracing, are associated with greater movement.

By virtue of this, we found a positive correlation between iconicity and movement, suggesting that more iconic signs tend to involve larger, more expressive movements.

2.6 Summary and motivations for the present study

To summarize, Zipf's (1949) Principle of Least Effort, which emphasizes the tendency for speakers to use the most efficient and least effortful forms of communication, has been further developed in later theories such as Grice's (1975) Maxim of Quantity, Sperber and Wilson's (1986) Relevance Theory, and then extended into discourse analysis through the work of Chafe (1976), Prince (1978), Gundel (1993), Givón (1983), and Ariel (1990), among many others. These theories converge on the simple idea that speakers tend to balance effort with communicative effectiveness, using minimal forms when context allows. In the context of sign languages, research has shown that similar principles apply: nominals are employed to mainly introduce referents and track lowly accessible referents, while zero forms such as CA and verbs track highly accessible referents but the findings on the role of classifiers w.r.t. accessibility is less conclusive (e.g., Barberà, 2013; Bel et al., 2015; Cormier et al., 2013; Czubek, 2017; Ferrara et al., 2022; Frederiksen & Mayberry, 2016; Morgan, 2000; Perniss & Özyürek, 2015; Pizzuto et al., 2006; Swabey, 2002). In this sense, establishing a cohesive discourse in sign languages mainly follows from the abovementioned theories and principles. Moreover, reference tracking has been studied more extensively in second language acquisition for both spoken (Ahrenholz, 2005; Hickmann & Hendriks, 1999; Williams, 1988) and sign languages (Bel et al., 2015; Frederiksen & Mayberry, 2016). It is known to be difficult for second language learners to acquire, given the complexity of the syntax-discourse interface (Lillo-Martin & Quadros, 2011). The general finding is that second language learners become more overexplicit in reference tracking (i.e., they produce more nominals and fewer null forms when tracking accessible referents) than natives. However, much less is known regarding the effects of delayed first language acquisition in the context of deaf adult signers on establishing a cohesive discourse, with the exception of a few studies (Becker, 2009; Gür, 2024; Gür & Sümer, 2022; Keleş et al., 2023).

Moreover, when proposing accessibility hierarchies for sign languages, most reference tracking studies used a binary (given vs. new) or tripartite (introduced, maintained, and re-introduced) classification of discourse context, and to my knowledge, there is only one work (Czubek, 2017) that quantized Ariel's accessibility by using Toole's (1996) scale to offer a more nuanced and granular analysis for ASL. Nonetheless, Czubek only included native signers (19 in total) in the experiment and did not explore the age of acquisition effects on the RE choice with increasing accessibility. Also, the stimuli used in the experiment mainly consisted of a simple wordless elicitation story with only six panels (images), which, as the author argued, might have contributed to the few occurrences of re-introductions in the participants' productions.

Additionally, through the recent advancements in computer vision, some studies showed how these principles of efficiency and phonetic reduction could be realized in sign languages, especially in the context of building a cohesive discourse. Using a markerless pose estimation methodology, Stamp, Dachkovsky, et al. (2024), for example, analyzed reduction patterns in ISL narratives and demonstrated that more discourse context predicts reduction. In other words, maintained nominal REs tend to undergo a reduction in duration and size, which is in line with the predictions of Ariel's accessibility hierarchy or Zipf's principle of least effort. Stamp, Dachkovsky, et al. appear to have included only native signers from a wide range of ages without specifying their age of acquisition status. Despite this, research suggests that age and age of acquisition might affect the extent of phonetic reduction, which implies that both factors need to be controlled.

These points are worth mentioning, as they point out an important gap in the existing body of literature. To my knowledge, there are only a few studies that look at the extent to which delayed first language acquisition influences cohesion or reference tracking strategies in sign language discourses and how it relates to phonetic reduction and referential accessibility. Another gap to highlight concerns the limited use of a storytelling paradigm to quantitatively measure accessibility scores in signed discourse across acquisition groups (native vs. late signers)and the lack of integration of kinematic or phonetic measures like duration, movement of the articulators, and size of the signing space to link these properties with referential accessibility. As a result, the primary motivation for this thesis is to bridge these areas by (i) employing a storytelling paradigm in TİD (a relatively understudied sign language) narratives with a quantitative measurement of referential accessibility scores based on the work of Ariel (1990) and Toole (1996); (ii) leveraging a markerless and open-source computer vision tool, MediaPipe, to detect phonetic reduction from 2D videos to investigate the relationship between discourse cohesion and phonetics, and (iii) examining the effects of delayed first language acquisition by comparing the discourse cohesion and phonetic strategies of deaf native and late signers.

CHAPTER 3

RESEARCH QUESTIONS AND METHODOLOGY

This chapter presents the research questions and describes the methodologies employed for discourse cohesion and phonetic analyses.

3.1 Research questions

This thesis examined the relationship among referential accessibility, chosen linguistic forms or referring expression (RE) types, and certain phonetic measures (e.g., duration, hand distance, and sign space use) across different acquisition groups (native and late signers) in Turkish Sign Language (TİD) narratives. Two types of analyses were conducted: analyses for discourse cohesion and phonetics analyses. The research questions for discourse cohesion are presented in 1 and 2 below.

(1) How do discourse context and type of referring expression relate to the accessibility scores of referents?

(2) Do native signers track accessible referents more than late signers, which is realized as minimizing the distance between current and previous mentions, tracking more topically salient referents, and engaging in less referential competition?

The data were collected as part of a larger project, funded by Boğaziçi University, Bilimsel Araştırma Projeleri (BAP), entitled "Supporting Sign Language Development of Deaf Children with Hearing Parents through Linguistically Informed Preschool Stories" (Start-up Project, #14458, PI: Kadir Gökgöz) and the analysis present in this thesis aims to build on the published papers from this project (Keleş & Gökgöz, 2022; Keleş, Atmaca, & Gökgöz, 2023).

A narrative production experiment was conducted with native (deaf-of-deaf) and late (deaf-of-hearing) adult signers of TİD in which participants watched ten clips, which were excerpts from the wordless cartoon Tom and Jerry, and they retold the events in the narratives. After the video-recorded elicitation, the narratives were annotated for RE Type (NOM, CL, CA, VERB) and Discourse (Introduced, Maintained, Re-introduced). A Python script was prepared to automatically assign referents an accessibility value on a 7-point scale (from -2 to 5). The scale is an adaptation of Toole's (1996) scale for English to TİD. The protocol evaluated each referent's distance from its previous mention, topicality/saliency in the discourse context, and the presence of competitors.

For the first question, it is predicted that TİD signers will be sensitive to the accessibility of referents. That is, they will use explicit forms with more phonetic content (e.g., full nominals) for referents with lower values of accessibility (i.e., the cases where using a non-economical language structure can be justified by the pragmatic necessities) and also in the introduced discourse context. However, I predict more implicit forms with less form (e.g., markers of zero anaphora and classifiers to some extent) for highly accessible referents, especially in maintained and re-introduced contexts.

I also expect an age of acquisition effect in RE choice. Our previous work (Keleş et al., 2023) suggested subtle differences between the native and late signers, especially in the production of implicit markers like zero anaphora. The late signers produced them to a lesser extent than natives, making them somehow over-explicit. Therefore, I predict that native signers might track referents that are more accessible referents than the late acquisition group. By the definition of the accessibility scale that I use in this thesis, tracking referents with a higher mean of accessibility will correspond to tracking the same referent over a period of time and then shifting as opposed to shifting back and forth between different and competing references.

The research questions for the phonetic analyses are presented 3 and 4 below. below.

(3) Do phonetic measures like sign duration, hand distance, and sign space use correlate with referential accessibility and discourse context in TİD narratives?

(4) Are such phonetic measures in the re-telling of events in signed narratives age-sensitive? In other words, does delayed or late exposure to first language affect sign duration, hand distance, and sign space use during the narration of an event in a story among deaf TİD signers?

A computer vision (CV) analysis using Google's Mediapipe library (Lugaresi et al., 2019) was used to address these questions. In the analysis, each 2D RGB frame across different narrative productions outputs 33 triplets (x, y, z) of the body pose landmarks for different locations on the signer's body to measure certain phonetic metrics like duration, hand distance, and sign space use (i.e., the Euclidean distance between wrist and a reference mid-shoulder point on the body). Out of the 33 extracted landmarks, I report 10 of them: left thumb, right thumb, left pinky, right pinky, left wrist, right wrist, left elbow, right elbow, left shoulder, and right shoulder.

To answer the first question on the interaction of cohesion and phonetics in reference tracking in narratives, I chose two REs: FARE 'mouse' and KEDI 'cat' (following a similar strategy employed in Stamp, Dachkovsky, et al., 2024). After marking the beginning and end of each of these lexical items, I calculated duration, Euclidean hand distance between each frame to estimate the amount of movement, and the distance between the wrist and the joint between the shoulders as a proxy to the size of the signing space. Each of these three measures was used as a dependent variable, and the effect of accessibility and discourse context were examined.

Considering the results of Stamp, Dachkovsky, et al.(2024), I predict that the accessibility value and the discourse context will also have a significant effect on the phonetic correlates of the referent. Namely, I predict the duration of the REs to be longer when a referent is introduced as opposed to when it is maintained or re-introduced. Furthermore, I hypothesize that movement amplitude and the signing space will be narrower for referents when they have higher accessibility values, and they will be larger for lowly accessible referents, supporting Zipf's principle of least effort (1949).

To investigate the second question, I conducted an utterance level analysis compared to a lexical/RE-based analysis. I input Mediapipe episodes from narratives that consist of one or more clauses and represent a particular event occurring in that particular story. The same phonetic measures were calculated using body pose landmarks.

To my knowledge, this is the first study that directly compares the phonetics/kinematics of narratives among deaf adult native and late signers. In a previous study with complex motion events used in isolation (Gökgöz & Keleş, 2023), we found that late signers of TİD used their non-dominant side of the body more than native signers. Other research, however, reported contrastive observations (Börstell, Schembri, & Crasborn, 2024; Braem, 1999) regarding the age of acquisition effects on duration. Thus, I predict that there might be a difference in the mean duration of utterances produced by native and late signers. In our previous study (Gökgöz & Keleş, 2023), a significant difference in the amount of movement only in isolated signs; I do not have a specific prediction for this measure. As for signing space use, Flaherty et al. (2023) also found sign space reduction for Nicaraguan Sign Language (NSL) among the acquisition cohorts, suggesting that this measure can be informative, especially for newly emerging sign languages. Since TİD is considered to be an established sign language, I do not predict a difference between the groups in the signing space use.

3.2 Methodology

3.2.1 Production experiment

This experiment employed a story-telling paradigm: The participants were shown ten different silent or wordless video clips, which were extracted from the cartoon Tom and Jerry, and then were prompted to retell what happened in each of the clips. The participants' reference tracking strategies were examined. The participants were all compensated for taking part in the experiment.

3.2.1.1 Participants

29 deaf adult signers participated in the experiment. 15 of the participants were native or deaf-of-deaf (DoD) signers who all had one or more deaf caregivers. All of the native participants self-reported that they started acquiring TİD from their TİD-signing parents before age three. The remaining 14 signers were classified as

late or deaf-of-hearing (DoH) signers. The criteria for being a late signer was to have hearing parents who were non-signing. The caregivers of the late signing participants were all non-signers, but due to the need for mutual communication, it is common for the household members to invent a gestural communication system called homesign. Homesign variations display internal consistency and structure to varying degrees; however, they are not treated as full-fledged natural languages.

The participants' background information is presented in Table 1. Factors such as years of education, number of male and female participants, the estimated number of years of TİD use, and other self-reported language measures for both TİD and Turkish productive and receptive skills) did not differ significantly for the two groups; *p*'s > .05. Ages, however, differed across the groups ($p < .01$). Natives (M_{age} = 26 years) were overall younger than late signers (M_{age} = 35 years). This was inevitable since we prioritized controlling for the estimated number of active TİD use.

The age of the native signers (seven males, seven females) ranged from 18 to 35 (M = 26 years, SD = 5 years). All of the native participants were prelingually deaf[6], except one native signer who acquired deafness after age four. The home language for all native participants included TİD. Three signers also reported the use of homesign. One participant (number 13) was excluded from the analyses because their background did not comply with the criteria for native classification. That is, they reported only using homesign (and not TİD) in the household during childhood as a primary means of communication with the parents, despite reporting having deaf caregivers. All signers completed primary and middle school in a deaf school in which TİD was used[7]. Furthermore, the number of years native participants attended formal education ranged from 12 (corresponding to the completion of high school) to

[6]Pre-lingual deafness refers to hearing loss occurring before language acquisition. re-lingual deaf participants were either born deaf or acquired deafness before age 3.

[7]According to the recent report by the Turkish Ministry of National Education, there are 72 deaf schools officially recognized in the country. TİD is primarily used in these schools. İlkbaşaran (2015) and Sarı (2005), among others, indicate that the students lack good quality instruction in sign language because almost all teachers are hearing and most even do not know how to sign although the attitude has started to change in recent years (based on anecdotal experience). As a result, most of the learning in deaf schools occurs in the form of peer learning.

Table 1. Participant demographics (reproduced from Keleş et. al., 2023)

No	Group	Gender	AoA	Age	LoE	Attended Deaf Schools	Home Language
2	Native	Male	0-3	33	12	PS, MS, HS	TİD, homesign
3	Native	Female	0-3	27	12	PS, MS, HS	TİD, homesign
13	Native	Female	0-3	27	12	PS, MS, HS	Homesign
14	Native	Male	0-3	29	12	PS, MS, HS	TİD, Homesign
18	Native	Male	0-3	28	12	PS, MS	TİD
20	Native	Male	0-3	32	12	PS, MS, HS	TİD
22	Native	Male	0-3	35	12	PS, MS, HS	TİD
24	Native	Female	0-3	23	12	PS, MS, HS	TİD
26	Native	Female	0-3	30	12	PS, MS, HS	TİD
27	Native	Female	0-3	25	12	PS, MS	TİD
29	Native	Female	0-3	24	14	PS, MS, HS	TİD
30	Native	Male	0-3	18	12	PS, MS, HS	TİD
33	Native	Female	0-3	21	16	PS, MS, HS	TİD
34	Native	Male	0-3	19	16	PS, MS, HS	TİD
38	Native	Female	0-3	25	12	PS, MS	TİD
10	Late	Female	4-7	35	14	PS, MS, HS	Homesign, TR
23	Late	Male	4-7	36	12	MS, HS	TR, homesign
4	Late	Male	4-7	33	12	PS, HS	TR
16	Late	Female	4-7	33	12	PS	TR, homesign
21	Late	Female	4-7	31	12	PS	Homesign
25	Late	Male	4-7	35	12	PS, MS, HS	Homesign
35	Late	Female	4-7	49	8	PS, MS	Homesign, TR
36	Late	Male	8-12	43	16	PS, MS	TR
12	Late	Male	8-12	28	16	PS	TR
15	Late	Female	8-12	24	12	PS, MS, HS	TR
19	Late	Female	8-12	24	12	PS, MS	TR
28	Late	Male	8-12	33	8	PS, MS	Homesign
31	Late	Male	13-17	50	8	MS, HS	TR, homesign
32	Late	Male	13-17	31	16	PS	TR

Note. PS = Primary School, MS = Middle School, HS = High School, TR = Turkish. In total, 38 participants attended the experiments conducted within the scope of the BAP project "Supporting Sign Language Development of Deaf Children with Hearing Parents through Linguistically Informed Preschool Stories" but not all completed the storytelling production experiment. For this reason, participant numbers reported here reflect only those who attended the production experiment as per my thesis.

16 years (corresponding to tertiary education). Two of the signers reported to have obtained a bachelor's degree, and one received an associate's degree.

The late signing group consisted of eight males and six females, with the ages ranging from 24 to 50 years (*M* = 35 years, *SD* = 8 years). Two of the late signers became deaf between the ages of four and seven, and the remaining participants were all prelingually deaf. Among the 14 late-signing participants, six reported learning TİD between the ages of four and seven. Six late signers began acquiring TİD between the ages of eight and twelve. The remaining two late signers reported to have learned TİD between the ages of 13 and 17. The primary language of communication in the household during the participants' childhood included mostly Turkish and homesign, with the exception of one signer (participant 10) who reported to have received TİD input from their hearing parents. I did not exclude this participant from the analyses because the TİD knowledge of the parents was only rudimentary, and their comprehension skills were poor, as reported by the participant. For this participant, the self-reported age of acquisition was nine years, and the estimated number of years for active TİD use (the onset was their self-reported age at which they started learning TİD and the offset was calculated as their current age) was 26 years, which is comparable to the native participants' TİD use. Like the native group, all late signers attended at least one school for the Deaf, in which a crude form of TİD was used by the teachers. The mean number of schooling years ranged between eight and 16 years, averaging 12 years (*SD* = 3 years). Two participants received an associate's degree, while three participants graduated from university and got a bachelor's degree. Three signers in the late signing group did not attend high school.

3.2.1.2 Video elicitation materials

Sample pictures from video clips are presented in Figure 2, and the video clips can be found at https://github.com/kelesonur/MA_Thesis. Each clip included 6 to 10 events, and the duration ranged from 17 seconds to 28 seconds (*M* = 23 seconds). The storylines consisted of events in which animate protagonists interacted with inanimate objects. Although the plot line was different for each stimulus, the context ensured the introduction, maintenance, and re-introduction of the tracked characters or objects in the clips. The detailed descriptions of the events in each video clip can be found in Keleş et al. (2023, pp. 29-31, Appendix B).

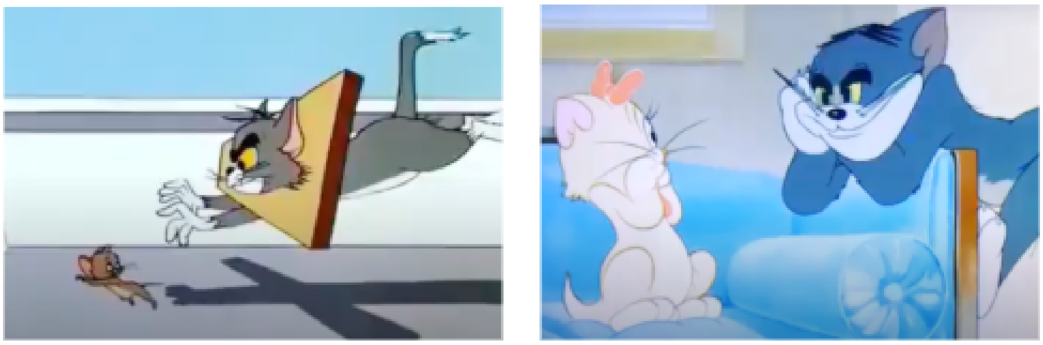


Figure 2. Sample pictures from video clips

3.2.1.3 Task procedure

Before the commencement of the experiment, each participant first filled out an informed consent form and then a background questionnaire on their deafness status as well as acquisitional and educational background. The contents of the informed consent and the background questionnaire were first translated into TİD and then explained by a research assistant. The informed consent form and the questionnaire can be found in Appendix A.

A research assistant or experimenter was present in all data collection sessions and gave detailed instructions for the task before the video recording began. A native deaf research assistant attended all the sessions, and in some sessions, she was accompanied by a hearing late TİD signer research assistant.

The participants took a seat in front of a laptop computer placed on a table and across from the assistant. The set-up of the experiment is exemplified in Figure 3. The visual stimuli were shown on the laptop screen. The session was recorded via an HD video camera, focusing on the participant's upper body. The participants were instructed that they were going to watch a clip and then retell what happened. The participants could watch the video as many times as they wished before narrating it. The participants watched the clips, provided a narrative, and proceeded to the next stimulus at their own pace by pressing a button on the keyboard. The participants were informed that they could take a short break during the experiment and that they had the right to leave the experiment for any reason.

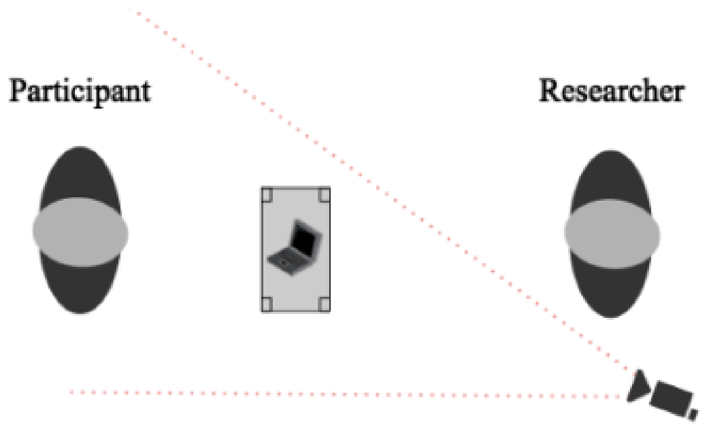


Figure 3. Experimental setting

The participants were told to narrate as if they were to one of their deaf friends. This was to ensure natural signing without feeling the need to adjust or modify their language to a non-native addressee. However, no clear instruction was provided about whether they should narrate to the hearing researcher directly or look into and narrate to the video camera. They were given this option because a native or native-like research assistant was not always available in the sessions. The majority of the signers (*N* = 25) preferred to narrate the clip directly to the camera in the absence of a native addressee, while the remaining four signers (two native, two late) completed the task by looking at the native addressee as they did not feel comfortable otherwise. However, this did not appear to affect the nature of the narratives with respect to referring expression (RE) use (see Appendix B for results of an analysis examining the main effect of interaction style (i.e., signers who narrated to the camera with those who did not).

3.2.1.4 Data annotation

The narratives were annotated by the author of this thesis and translated by a deaf bimodal-bilingual native research assistant into Turkish on ELAN Linguistic Annotation Software (Sloetjes & Wittenburg, 2008), following the guidelines in Berman and Slobin (2013) and Gullberg (2006). Clausal boundaries were determined based on the presence of predicates and certain prosodic or non-manual markers like a head nod. Table 2 includes the descriptions of the RE types.

Table 2. Description of clause boundary, discourse contexts, and RE types (reproduced from Keleş et. al., 2023)

Tier	Tag	Description
Clause Boundary		The boundary between clauses indicated by the presence of predicates and/or certain prosodic cues (e.g., head nod)
Discourse Context	Introduced	A first mention of an entity in a given discourse, regardless of any syntactic role
Discourse Context	Maintained	A subject referent continued across two or more immediate clauses
Discourse Context	Re-introduced	A subject referent discontinued across at least one clause
RE Type	NOM: Bare Noun	A non-modified noun phrase (e.g., CAT, MOUSE)
RE Type	NOM: Modified Noun	A noun phrase modified by an index sign, classifier, etc.
RE Type	NOM: Fingerspelled Noun	Borrowed nouns from Turkish spelled by the manual alphabet (e.g., K-E-D-İ 'cat')
RE Type	PRO: Pronominal IX	An index sign used in isolation to refer to an entity in discourse
RE Type	CL: Whole Entity Classifier	A handshape that refers to the entirety of an entity
RE Type	CL: Body Part Classifier	A handshape that refers to a part of an entity
RE Type	CL: Extension Classifier	A handshape that denotes the size and shape of an entity
RE Type	CA: Constructed Action	A multifunctional construction where the signers take on the role of a referent and perform their actions
RE Type	VERB: Plain Verb	A non-agreeing predicate (e.g., LOVE)
RE Type	VERB: Agreement Verb	A predicate the beginning and ending of which agree with the person and number of subject and object

Following the discourse context annotation criteria by Gullberg (2006), referents of any syntactic role (e.g., in the form of a subject or object) had the status introduced if they were brought to the discourse context for the first time in the narration of a particular stimulus. For example, if the signer introduced the referent Tom with a bare nominal KEDI 'cat' in the context of a new stimulus narrative, this referent received the introduced context, regardless of the referent being introduced by the same participant in an earlier narrative[8]. Subject referents were coded for maintenance and re-introduction. In other words, referents that were continued across two or more immediate clauses had the maintained context only if they were the grammatical subjects of the clause. In contrast, subject referents had the re-introduced context if it was discontinued across two or more immediate clauses or simply if there was at least one intervening clause between its previous and current mention that had a competing referent.

RE Type was categorized as nominal, pronominal, classifier, constructed action, and verbal. The nominal category branches into modified nouns (e.g., nouns that are followed or preceded by classifiers or index signs), bare nouns, and fingerspelled nouns, which refer to the borrowing from Turkish and the spelling of nouns using the manual alphabet (e.g., F-A-R-E 'mouse). Pronominal forms are the isolated uses of the index sign used to make a reference to an entity. Classifiers (CL) were annotated as whole entity classifiers (WECL), body part classifiers (BPCL), as well as tracing and static size and shape specifiers, which are handshapes that can refer to the entirety, part, and size and shape of an entity. Given their similar function of size and shape denotation, tracing and static specifiers were collapsed together with WECL and BPCL. Constructed action (CA) occurs when signers embody character referents, using manual (e.g., limbs) and non-manuals (e.g., facial expressions) to act out actions and feelings. Finally, verbal categories (VERB) were annotated as plain or agreement verbs depending on the type of the predicate in the

[8]The accessibility scores, the calculation of which is explained in the following subsection, for these two introduced referents will be different, with the second instance having a slightly larger accessibility score.

clause. The separation of CA from VERB followed previous literature. Differently from verbal constructions, CA spreads over the entire clause and often necessitates the activation of both non-manual and manual features (which involves the utilization of most body parts). In addition, CA can track multiple referents, one from the character's perspective and others from the observer's perspective.

The example in (5) illustrates how the REs were annotated. The signer starts a new clause whose subject (the male cat) was tracked in the immediately previous context; therefore, the glosses WALK.CA and POSE.CA, representing the whole clause with signed constructed action, were regarded as the maintained RE for the male cat, although the nominal anaphor is omitted. In the next clause, the signer explicitly introduces a new referent (fish aquarium) into the discourse by signing a noun compound. This referent acts as the indirect object of the sentence. Since referents with introduced contexts can have different syntactic roles in the sentence, the annotation of the fish aquarium is justified. The subject referent in the sentence (the male cat) is, however, again omitted and tracked with an agreement verb (give). Its discourse context was annotated as maintained because of the mention of the subject referent (the male cat) in the preceding clause. However, the direct internal argument of the verb in the sentence (her or female cat) was not annotated since the referent is not introduced for the first time and is not the subject.

In case there were multiple REs employed concurrently, I only annotated the expression that tracked the primary subject referent. The example in (5) illustrates this. In (5), the last frame depicts two perspectives, by means of which the signer tracks two referents at the same time. The signer's non-manuals (e.g., facial expressions and head position) tracked the actions of the male cat through CA, whereas the manual handling classifier (signed with the right hand) simultaneously represented the fish. Since the second referent (i.e., the fish) is not the subject referent, it has not been annotated.

(5) 

WALK:CA POSE:CA FISH AQUARIUM $_{1}$GIVE$_{3}$

'(The male cat) walks and comes near (the female cat). (He) offers (her) a fish (aquarium).'

A second coder also watched the clips and divided each clip into meaningful events. There was a strong correlation (r = .85) between the annotations by the two coders. I segmented the plot in each of the clips into different events (minimum: 6, maximum: 10, on average: 8 events). I then annotated the start and end points of each event for each participant. The rater later determined the start and end of each event. Events, the start and end of which only differed by five or fewer video frames[9] were chosen for the computer vision analysis. If an event was skipped or partially told (missing one or more details) in a narrative, it was coded as such. Additionally, points were given based on reported events for each narrative clip. All explicit mentions of events received a score of 1, incomplete mentions received a score of 0.5 (e.g., some parts of the events were missing), and missed/skipped events received 0. More detailed information on event analysis is provided later in this chapter.

Following the manual annotations, the ELAN files were exported to a comma-separated value (.csv) file with an addition of the participant number, acquisition group, clip number, event number, points (assigned to clips based on completed, partially completed, or missing events), referent name, discourse context, and RE type. The .csv data were further analyzed for cohesion and phonetic analysis.

3.2.1.5 Assigning accessibility scores

Following the manual annotations on ELAN, each RE was also assigned an accessibility score using Toole's (1996, p. 273) accessibility rating scale, which was originally developed for English across four different genres (see Czubek, 2017 for

[9]A video frame is an individual, static image within a sequence of images that are rapidly displayed to generate the perception of motion in video media.

its exact implementation in American Sign Language - ASL). I adapted this scale to TİD for the purposes of this study. Toole's scale quantifies the potential difficulty of retrieving an entity within a given discourse context on a scale of 6 (most accessible) to -2 (least accessible). The calculation of the accessibility score is based on three pivotal factors: distance and unity, topicality, and competition. My adaptation of this protocol differs from the original only in the scoring of the distance and unity algorithm and, therefore, ranges from 5 (most accessible) to -2 (least accessible). One reason for this is that I focus on subject REs for maintenance and re-introduction, following the annotation criteria in the reference tracking literature (e.g., Frederiksen & Mayberry, 2016, 2019; Gullberg, 2006). Furthermore, the original scale used the term "episode" to mean a set of propositions that compose a semantic unit in a discourse, similar to a paragraph in a written composition. Since the signers in this study narrated clips that were connected but independent (e.g., they have different plots but share the same protagonists), I took each clip narrative to be one episode and rated the unity factor accordingly. The algorithm for each of the factors was as follows:

Distance and Unity: The scoring of distance was determined by the proximity of the last mention and whether referents share the same narrative context or not. This score decreased as the distance in textual context between the current and previous mentions increased, reflecting the cognitive load required to recall and track the referent, such as the following:

For a referent R at point t in the discourse (R_t):

a. If the last mention of R is in the clause immediately preceding that of R_t, and they are both in the same narrative context, then the accessibility of R_t is 3.
b. If the last mention of R is in the same narrative context with R_t, but not in the same or preceding clause, then the accessibility of R_t is 2.
c. If the last mention of R is in the narrative context immediately preceding that of R_t, then the accessibility of R_t is 1.

d. Otherwise, the accessibility of R_t is 0.

Topicality: Topicality was evaluated based on how frequently the referent appeared in the discourse context, specifically within the last four clauses prior to the current mention. More frequent mentions increased the score, enhancing the entity's salience and, hence, its accessibility. This increase was by one point for one or two mentions and two points for more than two mentions, as in the following:

For a referent R at point t in the discourse (R_t):

a. If R has been mentioned more than twice in the last four clauses relative to R_t, then the accessibility of R_t is increased by 2.

b. If R has been mentioned once or twice in the last four clauses relative to R_t, then the accessibility of R_t is increased by 1.

c. Otherwise, the accessibility of R_t does not change.

Competition: Competition scores were related to the presence of competing entities that might confuse or dilute the listener's or reader's ability to recall the referent to be tracked. If other referents were mentioned between the last and the current mention of the focal referent, the score was adjusted downwards. This adjustment was made by one point for one competing entity and two points for multiple competitors, reflecting the increased effort in distinguishing between multiple entities. That is:

For a referent R at point t in the discourse (R_t):

a. If there are two or more competing referents between R_t and the last mention of R, then the accessibility of R_t is decreased by 2.

b. If there is one competing referent between R_t and the last mention of R, then the accessibility for R_t is decreased by 1.

c. Otherwise, the accessibility of R_t does not change.

Note that the calculation of accessibility scores was nested such that each of the abovementioned factors (i.e., distance and unity, topicality, and competition) was applied in the sequence presented above. The base accessibility score for a referent was initialized based on its proximity from its previous mention. Then, the score was incremented according to its topicality. Finally, the score was decremented based on the number of competing entities.

The examples in (6,7,8) provide an illustration of how the accessibility score was calculated for each referent or entity. Each example belongs to the same narrative and represents the first three consecutive clauses from a signer's first narrative. All examples contain tiers for gloss, the referent's syntactic role, its discourse context and type, scores for three primary factors (distance and unity, topicality, and competition), and the total accessibility score.

(6) Clause 1


CAT


MOUSE


CHASE

Referent's Syntactic Role:	Subject	Object
Discourse Context:	Introduction	Introduction
RE Type:	Nominal	Nominal
Distance and Unity Score:	0	0
Topicality Score:	0	0
Competition Score:	0	0
Total Accessibility Score:	0	0

Translation: 'Tom is chasing Jerry.'

In (6), the signer tracks two referents, Tom and Jerry, by signing the nominals for CAT (functioning as the subject) and MOUSE (functioning as the object). The discourse context indicates that both referents are being introduced for the first time in the narrative; in other words, they lack prior mentions. In terms of accessibility scoring, both produced referents each begin with a distance and unity score of 0, as they have not been mentioned previously, reflecting their status as new entities. However, if this is not the first narrative context and the signer has already tracked these two referents in the previous narrative context (same characters but in a different plot), then both referents have the distance and unity score of 1. Moving to topicality, neither CAT nor MOUSE appear in prior clauses, so both receive a topicality score of 0. According to this score, neither referent has become more salient by being mentioned more than once in the last four clauses, which would increase their accessibility. Competition is the final factor to consider. In this case, CAT and MOUSE (both mentioned for the first time) do not face any competing referents, as there are no other entities requiring attention, so their competition score remains 0. The total accessibility score is then calculated by summing all these factors, resulting in an accessibility score of 0 for CAT and MOUSE.

(7) Clause 2


TUNNEL


JUMP

Referent's Syntactic Role:	Adjunct	Subject
Discourse Context:	Introduction	Maintenance
RE Type:	Classifier	Classifier (Right Hand)
Distance and Unity Score:	0	3
Topicality Score:	0	1
Competition Score:	-2	-1
Total Accessibility Score:	-2	3

Translation: 'Jerry is jumping into a tunnel.'

The example in (7) features two referents: Tunnel, serving as the adjunct, and Jerry, which is the null subject expressed by the predicate sign JUMP. The discourse context introduces the sign TUNNEL while maintaining the continuity of Jerry as the subject, whose referent has already been established in the previous clause. Classifiers are used for both referents. As for distance and unity, Tunnel receives a score of 0, as it is introduced for the first time in the discourse and lacks any prior mention. On the other hand, Jerry is assigned a score of 3, as it was mentioned in the immediately preceding clause, maintaining the same narrative context, thus making it highly accessible. Recall that maintenance and re-introduction contexts are only considered for grammatical subjects, whereas introduced referents are annotated regardless of the syntactic role of a discourse referent. For topicality, Tunnel receives a score of 0 because it is new, i.e., an introduction because it has not appeared in the last four clauses. In contrast, Jerry has already been mentioned once in the discourse, resulting in a topicality score of 1. For the competition, Tunnel has two competing entities from prior clauses (the signs CAT and MOUSE tracking the referents for Tom and Jerry), leading to a decrease of 2 points, while Jerry faces only one competing referent (Tunnel) between its current and previous mention, so its score is decreased by 1 point. Therefore, the final accessibility scores are -2 for Tunnel and 3 for Jerry.

Finally, let's consider (8), which again involves the referents CAT and MOUSE, both functioning as subjects within the sentence. The discourse context signals the re-introduction of both referents (both expressed nominally), indicating that they have appeared before but are being brought back into the context. The distance and unity score for both CAT and MOUSE is 2, as they were mentioned earlier in the discourse context but not in the immediately preceding clause.

(8) Clause 3

CAT MOUSE CRAWL.CA

Referent's Syntactic Role:	Subject	Subject
Discourse Context:	Re-introduction	Re-introduction
RE Type:	Nominal	Nominal (Right Hand)
Distance and Unity Score:	2	2
Topicality Score:	1	1
Competition Score:	-2	-1
Total Accessibility Score:	1	2

Translation: 'Tom and Jerry start crawling (in the tunnel).'

In terms of topicality, both CAT and MOUSE appeared once or twice in the last four propositions, thus each receiving a score of 1. However, competition again plays a role in decreasing the accessibility of both referents. CAT is assigned a competition score of -2, as there are two competing referents (TUNNEL and CAT) between its last and current mention. Meanwhile, MOUSE faces only one competing referent, resulting in a competition score of -1. The total accessibility scores are calculated as 1 for CAT and 2 for MOUSE, reflecting that both have moderate accessibility, though MOUSE is slightly more accessible due to facing less competition.

This accessibility scoring protocol was automated using a rule-based Python (3.12.4) script. The program inputted the .csv file obtained after the ELAN annotations and outputted the accessibility scores for each RE in a separate column. The data and script can be found here:
https://github.com/kelesonur/MA_Thesis/tree/main/Assigning_Accessibility_Scores

3.2.2 Phonetic analysis

To investigate the phonetic component of narrative production in TİD, I used a pose estimation library. I proposed two analyses: a RE analysis (on a lexical level) and an event analysis (on an utterance level). While the RE analysis examines the relationship between phonetics and discourse cohesion in signed narratives, the event or utterance-level analysis aims to reveal age of acquisition (AoA) effects among native and late TİD signers. I report the following phonetic measures: (i) sign duration, (ii) accumulated hand distance, and (iii) signing space size.

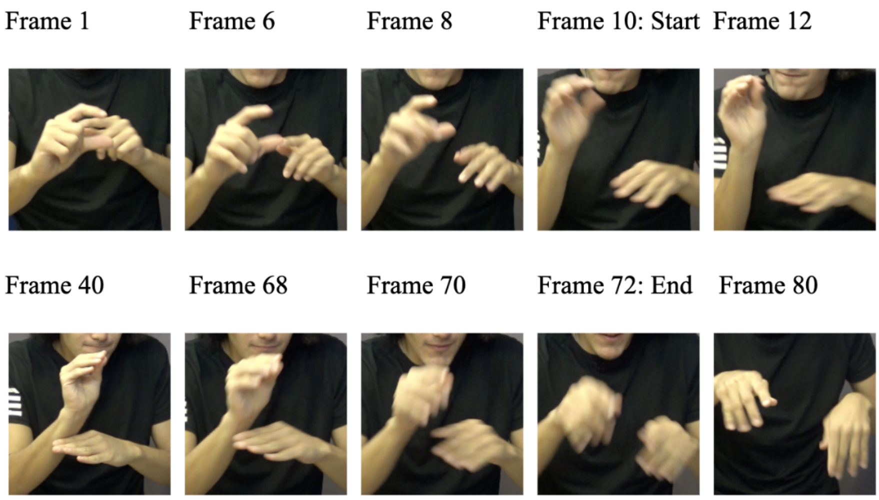
3.2.2.1 RE analysis

For this analysis, I chose two lexical items or bare nouns from the signer productions obtained in the first experiment: FARE ‘mouse/Jerry’ and CAT ‘cat/Tom’. Other RE types were not considered for analysis, and there are several reasons for this decision. Some RE types, such as constructed action, do not have a clear onset and offset boundary marked by certain prosodic cues, and it is often difficult to understand the scope of such constructions. While the beginning and end points of some other tools, like classifiers, can be more accurately marked, different classifier forms can be employed to track the same referent since the use of the classifier can change depending on the narrative context. For example, Tom can be represented with a body part classifier when the context has him in a sitting position, but the same referent can be tracked with a whole entity classifier with a 1-handshape (upright being) if he is standing in a different context. Given this variation and lack of systematicity in other forms, I decided to analyze only bare nouns, which do not vary from one discourse context to another.

Following previous work (Stamp, Dachkovsky, Hel-Or, Cohe, & Sandler, 2024), a deaf native TİD signer, who was not informed about the methodology and research aims of this thesis, identified the first frame as the frame in which the signer's hand stopped contact and moved from the rest position (e.g., the signer's legs), or the handshape and orientation for the sign became identifiable. Accordingly, the last frame of each bare noun was identified as the frame where the signer's hands returned to rest, or there was an identifiable switch in the handshape and orientation.

The example in (9) illustrates the identification of the first and last frame of the target lexical item FARE 'Mouse', which includes the frames from 10 to 72 in the given example. Frame 1 shows the sign that immediately precedes FARE, which is a two-handed asymmetrical sign. The handshape required for the target sign FARE is a "narrowed o". The "c-handshape" of the previous sign can be considered a static specifier, representing a tunnel. Frame 6 depicts the start of the transition from the previous sign into the target sign FARE, as can be deduced by the increasing distance between the fingers. Frame 10 is considered to mark the start of FARE since the handshape required for the sign, i.e., a "narrowed o", and the orientation becomes identifiable although the sign has not yet reached its complete form. Frame 40 represents the full lexical item when all the transitions are completed. Frames 68-72 are the last frames of the sign where there is a clear loss of contact between the dominant (i.e., right) hand and the nondominant (i.e., left) hand. Frame 72 demonstrates the selected final frame for the sign FARE again, owing to the identifiable switch in the handshape and orientation of both hands. This time, the handshape changes from that of the target sign FARE to the handshape of the next sign.

(9)



3.2.2.2 Event analysis

In addition to a lexical analysis, an utterance level analysis was conducted to investigate the AoA effects among native and late signers. For this, nine specific events were selected from various narratives as a unit of analysis out of the possible 80 events in total (8, on average, per narration clip), resulting in 233 data points in total produced by all the participants. Since I previously annotated which events were skipped and which were partially narrated, this information was used as the criteria for determining the best 'remembered' events. It would not make sense to compare a small event that was ignored or deemed not worthy of narrating by the participants with a striking event that was remembered and narrated by every participant. The criteria I followed are as follows: A score of 1 was assigned to all explicit mentions of events, a 0.5 to incomplete mentions (e.g., some parts of the events were missing), and a 0 to missed/skipped events. As a non-native TİD signer, I annotated the start and end points of all the events in the data. Furthermore, a second annotator, who was not informed about the experimental design (also a non-native TİD signer), also marked the start and end point of each event, and I analyzed the time difference (in ms) between two sets of annotations. The event selection was based on two criteria:

(i) how well the events were remembered by the participant and (ii) agreement on the start and end frames of events between the two annotators. I filtered independent events that received a score of above 0.9 (i.e., all or most of the details were remembered and reproduced well by the participants). This resulted in 25 events (out of 80 possible events across all narrative clips). Out of these 25 events, I chose the ones in which both annotators differed by only ten frames or fewer when marking the start and end timestamps for the events. The reliability subset included events that had an average length of 278 frames (all videos have 50 frames per second, and the average event duration is 5.56 seconds). By virtue of this, an event's first and last frame identification agreement percentage ranges from 97% to 99% (cf. Flaherty et al., 2023 for a similar analysis). This totaled up to 9 unique events per participant. The events chosen for this utterance-level analysis are as follows:

1. Event 5 - Clip 2: Score: 97%; Interrater Frame Dif: 4 (start) and 2 (end) frames.
2. Event 6 - Clip 2: Score: 95%; Interrater Frame Dif: 3 (start) and 6 (end) frames.
3. Event 7 - Clip 2: Score: 92%; Interrater Frame Dif.: 6 (start) and 1 (end) frames.
4. Event 2 - Clip 3: Score: 95%; Interrater Frame Dif.: 7 (start) and 1 (end) frames.
5. Event 4 - Clip 3: Score: 92%; Interrater Frame Dif.: 2 (start) and 1 (end) frames.
6. Event 9 - Clip 4: Score: 98%; Interrater Frame Dif.: 8 (start) and 1 (end) frames.
7. Event 4 - Clip 6: Score: 95%; Interrater Frame Dif.: 7 (start) and 1 (end) frames.
8. Event 8 - Clip 6: Score: 97%; Interrater Frame Dif.: 9 (start) and 2 (end) frames.
9. Event 5 - Clip 10: Score: 100%; Interrater Frame Dif.: 1 frame (start and end).

CHAPTER 4

DISCOURSE COHESION ANALYSIS

This chapter presents the analyses and results for discourse cohesion in TİD.

4.1 Data analysis

Pronominals were excluded from the analyses as their use was very infrequent (40 in total, amounting to 1% of the data). The remaining four RE type levels used in the statistical analyses were NOM, CL, CA, and VERB. The outcome variable (Accessibility Score) was treated as a continuous variable. Sum contrasts (effects coding) were used for the categorical variables in all the models. This coding scheme compares each level to the grand mean of the factor, and the coefficients sum to zero (Brehm & Alday, 2022). Sum contrast was used in this study because there was no natural reference or baseline category among the eight RE type levels. Thus, using sum contrasts allowed me to treat all categories equally, ensuring that each level's effect was compared to the overall grand mean rather than to an arbitrary reference group.

During data annotation, I noticed a native signer (participant number 34) who extensively used zero anaphora (more specifically, Constructed Action or CA, 77.9% compared to $M = 24.2\%$) and employed very few nominals (*M* = 6.2%) compared to the mean of the whole group ($M = 49.2\%$). Data from that participant was, therefore, excluded from the main analyses (Please see Appendix C for details).

Accessibility scores were analyzed with mixed effects regression models using the `lmer` function of the `lme4` package, version 1.1-27 (Bates et al., 2024) (see Czubek's 2017 PhD dissertation, where he also treated accessibility scores as a continuous dependent variable). In the models, RE Type, Discourse, and Group were the predictor variables. However, RE type and Discourse were never entered as predictors in the same model. The rationale behind this choice is twofold. Firstly, the model that included both RE Type and Discourse was not significantly better than the model that only contained Discourse as the predictor, $\chi^2(7) = 13.78, p = .055$.

Furthermore, to investigate the multicollinearity between these two predictors, a variance inflation factor (VIF) was calculated using the `vif` function of the `car` package 3.0-10 (Fox, Friendly, & Weisberg, 2013). The VIF score between RE Type and Discourse indicated slight multicollinearity (2.50), and it was twice as great as the VIF score between Group and RE Type (1.00) as well as Group and Discourse (1.00), both of which indicated zero multicollinearity. As a result, Section 4.2.1 reports on the descriptive and inferential findings for the effects of Discourse and Group on Accessibility. Section 4.2.2 reports the descriptive results as well as the model specifics, model comparisons, and inferential results for the effects of RE Type and Group on Accessibility Score.

The models with Discourse as a predictor had Group (Native, Late) and Discourse (Introduction, Maintenance, and Re-introduction) as fixed effects and participant and clip number as random effects.

In the models with RE Type as predictor, Group (Native, Late) and RE Type (NOM, CL, CA, VERB) were entered as fixed effects, and participant and clip number were random effects.

It is important to note that while these analyses modeled accessibility scores as the dependent variable influenced by Discourse and RE Type, we recognize that, theoretically, the accessibility of a referent can be viewed as an inherent property that affects such factors. However, initial empirical observations indicated that different RE types were associated with varying accessibility scores across groups and discourse contexts. Modeling accessibility as the dependent variable allowed us to explore these associations through pairwise comparisons (e.g., how the accessibility scores of each form differed from one another). These comparisons then allowed us to establish an accessibility hierarchy of RE forms, which is discussed in detail in Chapter 6. However, to align our analysis with theoretical considerations, we conducted an additional multinomial logistic regression analysis, treating RE Type (with four levels) as the dependent variable predicted by other factors, which can be found in Appendix D[10].

4.2 Results

The data were imported into the R statistical programming language, version 4.1.0, for statistical analyses. As mentioned above, data from one participant (P34) was excluded from the analyses, resulting in a total of 4069 data points. I will first report the descriptive statistics for the percent RE use for introduced, re-introduced, and maintained discourse contexts across native and late signing groups. I will then present the data on how RE Type, Group, and Discourse context affected the accessibility scores.

Figure 4 presents the distribution of RE Type across different Discourse Contexts (Introduced, Maintained, Re-introduced) and Group: (a) Native and (b) Late.

Out of the 4069 produced REs, 1057 (26%) of them introduced a referent, 1638 (40%) were used to maintain the previous referent, and 1374 (34%) were used for re-introduction.

Both natives and late signers used nominals as the primary strategy to introduce referents (i.e., the first mention of a referent) ($M = 94.1\%$, $M_{\text{Native}} = 95.2\%$, $M_{\text{Late}} = 93.2\%$), which was followed up by tracing size and shape specifiers ($M = 3.4\%$, $M_{\text{Native}} = 2.6\%$, $M_{\text{Late}} = 4.1\%$) that were also often used to introduce referents into the discourse context and they were used relatively more by late participants than by native participants. Other RE types, such as pronouns ($M = 0.4\%$, $M_{\text{Native}} = 0.2\%$, $M_{\text{Late}} = 0.5\%$), classifiers (WECL: $M = 0.6\%$, $M_{\text{Native}} = 0.8\%$, $M_{\text{Late}} = 0.4\%$; BPCL: $M = 0.3\%$, $M_{\text{Native}} = 0.4\%$, $M_{\text{Late}} = 0.2\%$), CA ($M = 0.2\%$, $M_{\text{Native}} = 0.0\%$, $M_{\text{Late}} = 0.4\%$), and verbal constructions (PLAIN: $M = 0.1\%$, $M_{\text{Native}} = 0.0\%$, $M_{\text{Late}} = 0.2\%$; AGR: $M = 0.1\%$, $M_{\text{Native}} = 0.2\%$, $M_{\text{Late}} = 0.0\%$) were minimally used to introduce context.

[10] The results of the multinomial analysis supported the results of the linear regression models reported in this thesis, indicating a high correlation.

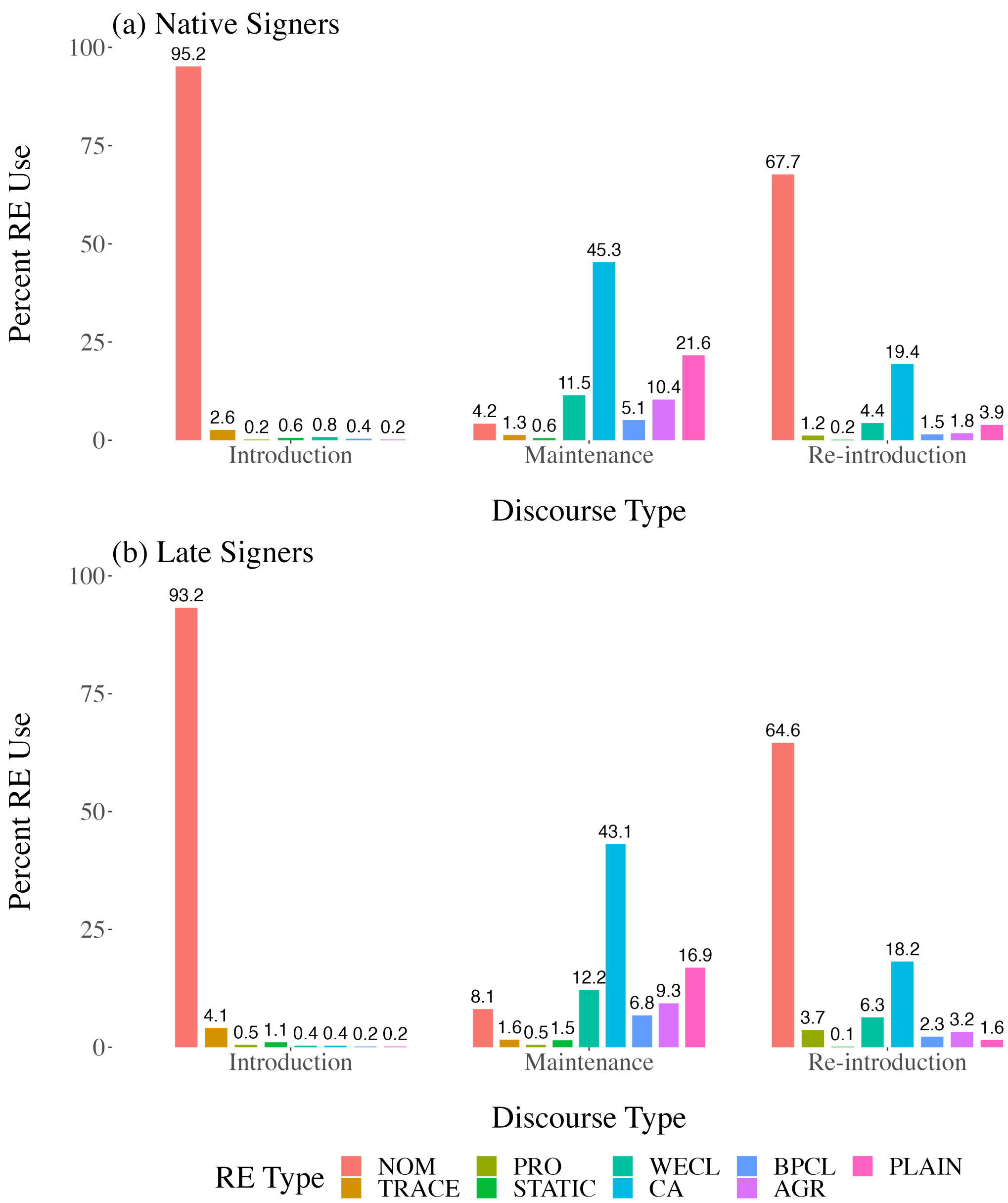


Figure 4. Percent RE Type Use by Discourse for native (a) and late signers (b). Note. NOM = Nominal; TRACE = Tracing Size and Shape Specifier; PRO = Pronoun; STATIC = Static Size and Shape Specifier; WECL = Whole Entity Classifier; CA = Constructed Action; BPCL = Body Part Classifier; AGR = Agreement Verb; PLAIN = Plain Verb

CA and verbs were the most commonly employed REs to maintain a previously established referent, accounting for approximately 80% of all constructions. Among these markers, CA was the most preferred tool, more so among the native signers than the late group ($M = 44.3\%$, $M_{\text{Native}} = 45.3\%$, $M_{\text{Late}} = 43.1\%$). Plain verbs were the second most utilized constructions in the maintained context ($M = 19.5\%$, $M_{\text{Native}} = 21.6\%$, $M_{\text{Late}} = 16.9\%$), followed by agreement verbs ($M = 9.9\%$, $M_{\text{Native}} = 10.4\%$, $M_{\text{Late}} = 9.3\%$). Moreover, in this discourse context, there was a notable decline in the use of nominals ($M = 6.0\%$, $M_{\text{Native}} = 4.2\%$, $M_{\text{Late}} = 8.1\%$). Overall, late signers used almost twice as many nominals to maintain a previously established referent as natives. Verbs and CA were followed by classifiers for maintenance. WECL occurrences ($M = 11.8\%$, $M_{\text{Native}} = 11.5\%$, $M_{\text{Late}} = 12.2\%$) doubled BPCL ($M = 5.9\%$, $M_{\text{Native}} = 5.1\%$, $M_{\text{Late}} = 6.8\%$) across all signers. The other RE types, including pronouns ($M = 0.2\%$, $M_{\text{Native}} = 0.0\%$, $M_{\text{Late}} = 0.5\%$) and size and shape specifiers (TRACE: $M = 1.5\%$, $M_{\text{Native}} = 1.3\%$, $M_{\text{Late}} = 1.6\%$; STATIC: $M = 1.0\%$, $M_{\text{Native}} = 0.6\%$, $M_{\text{Late}} = 1.5\%$) all had very low frequency.

To re-introduce a referent, nominals again were the prevalent tools ($M = 64.1\%$, $M_{\text{Native}} = 67.7\%$, $M_{\text{Late}} = 64.6\%$), and they were followed by CA ($M = 18.8\%$, $M_{\text{Native}} = 19.4\%$, $M_{\text{Late}} = 18.2\%$). When we look at the percent RE use, like in maintenance, native signers produced more instances of null forms, including the plain verb category ($M = 2.7\%$, $M_{\text{Native}} = 3.9\%$, $M_{\text{Late}} = 1.6\%$) but excluding the agreement verbs ($M = 2.5\%$, $M_{\text{Native}} = 1.8\%$, $M_{\text{Late}} = 3.2\%$). As opposed to reference maintenance, classifier use was higher than verbal constructions when signers brought a referent back into discourse. Similar to other contexts, WECLs ($M = 5.4\%$, $M_{\text{Native}} = 4.4\%$, $M_{\text{Late}} = 6.3\%$) were used considerably more often than BPCL instances ($M = 1.9\%$, $M_{\text{Native}} = 1.5\%$, $M_{\text{Late}} = 2.3\%$) for re-introduction. One important observation is that TİD signers used pronouns very sparsely but almost exclusively for referent re-introduction ($M = 2.5\%$, $M_{\text{Native}} = 1.2\%$, $M_{\text{Late}} = 3.7\%$). Late signers also used pronominals more than natives. The usage of other RE types for re-introduced context, such as size and shape specifiers (TRACE: $M = 0.0\%$; STATIC: $M = 0.1\%$, $M_{\text{Native}} = 0.1\%$, $M_{\text{Late}} = 0.2\%$), was either non-existent or very minimal, occurring below 1% of the time.

In Keleş et al. (2023), we had analyzed the same count data with a multinomial mixed effects regression with the `brms` package (Bürkner, 2019). In that analysis, RE Type was the dependent variable, and it had the following levels: Nominal, Classifier (STATIC, TRACE, WECL, BPCL), and Zero Anaphora (CA, PLAIN, AGR). Both Discourse and Group were the fixed effects in the model. We had found that Discourse significantly affected RE usage. Specifically, in the introduction context, nominal references were more likely ($\beta = 3.72$, CI: 3.34 – 4.13), while the maintenance context significantly decreased the likelihood of nominal REs ($\beta = -5.34$, CI: -5.70 – -4.98). On a similar ground, zero anaphora (i.e., *CA* together with verbal constructions) was less frequent in introduction contexts ($\beta = -3.87$, CI: -3.13 – -4.77), and, conversely, maintenance led to an increase in zero anaphora use ($\beta = 2.17$, CI: 1.75 – 2.66). Moreover, we found that being a native signer (compared to a late signer) influenced the RE use: There was no significant difference between natives and late signers in the use of nominals ($\beta = 0.06$, CI: -0.31 – 0.55). On the other hand, native signers were significantly more likely to use zero anaphora compared to late signers ($\beta = 0.89$, CI: 0.45 – 1.40).

Furthermore, the interaction between Group and Discourse in Keleş et al. (2023) also revealed that being a native signer influenced RE use, particularly in the introduction context, where tracking references with zero anaphora was more likely ($\beta = 1.58$, CI: 0.10 – 3.32) compared to late signers.

In this thesis, I did not reanalyze the count data, but rather, I was interested in how accessibility scores were influenced by these three factors (i.e., RE Type, Discourse, Group). Thus, the following subsection reports on the distribution of accessibility scores across RE Type, Discourse, and Group, and presents inferential results.

4.2.1 Effects of discourse and group on accessibility score

Figure 5 summarizes the data for accessibility scores (ranging between -2 and 5) by discourse context across all signers.

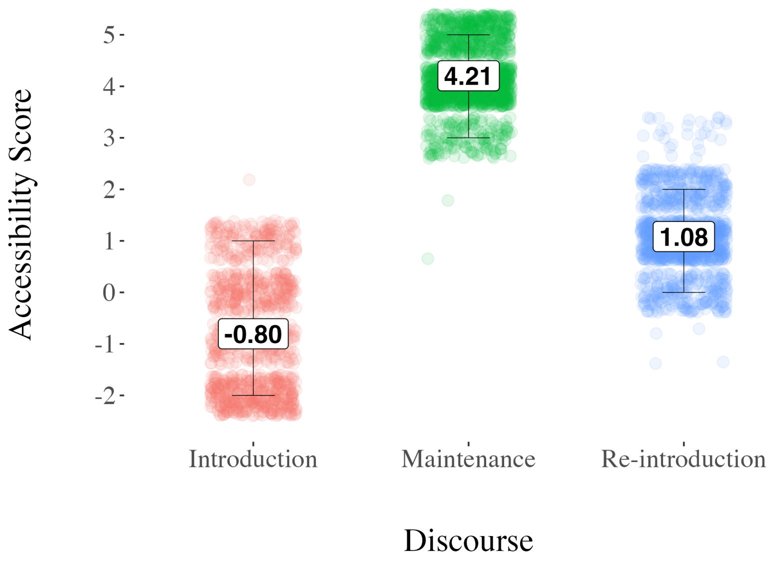


Figure 5. Mean Accessibility Score by Discourse Context. The text label represents the mean accessibility score, and the error bars represent the range between the 10th and 90th percentiles

As Figure 5 shows, accessibility scores also align well with discourse context. That is, introduced referents have low accessibility ($M_{\text{Accessibility}} = -0.80$), followed by the re-introduced referents ($M_{\text{Accessibility}} = 1.08$), and the maintained referents have the highest accessibility scores ($M_{\text{Accessibility}} = 4.2$).

Multiple mixed-effects models were fitted with Discourse and Group as predictors. Models were fit, ranging from a simple one-predictor model to a two-way interaction model. The model with the two predictors (Group, Discourse), as shown in (30), was significantly better than the model with only one predictor (Discourse), $\chi^2(1) = 9.90, p < .002$. However, the discourse model that included a two-way interaction between Group and Discourse was not a better fit, $\chi^2(2) = 0.42$, $p = .807$. Therefore, the results of the model in (1) will be reported below.

(1) Model_1 = lmer(Accessibility   Group + Discourse + (1 | Participant) + (1 | Stimuli))

The results of the model are presented in Table 3. Similar to the first model, sum contrasts were used, and the estimate for one level of each categorical variable (Re-introduction for Discourse and Late for Group) was not included in the model output. The results indicated that being a native signer significantly increased accessibility scores ($\beta = 0.08$, CI: 0.03 – 0.13, $p = .001$), as in the first model. Moreover, the discourse context significantly influenced accessibility scores. In particular, introduction was associated with substantially lower accessibility scores ($\beta = -2.28$, CI: -2.31 – -2.24, $p < .001$), while the maintenance context greatly increased accessibility ($\beta = 2.68$, CI: 2.65 – 2.71). This supports the observation that introduced referents have the lowest accessibility, re-introduced referents have moderate accessibility, and maintained referents have the highest accessibility.

Table 3. Model results for Accessibility Score by Group and Discourse

	Accessibility Score		
Predictors	Estimates	CI	p
Intercept	1.51	1.40 – 1.62	<0.001
Group [Native]	0.08	0.03 – 0.13	0.001
Discourse [Introduction]	-2.28	-2.31 – -2.24	<0.001
Discourse [Maintenance]	2.68	2.65 – 2.71	<0.001
Random Effects			
σ^2	0.60		
τ_{00} Participant	0.01		
τ_{00} Clip Number	0.03		
ICC	0.06		
N Participant	28		
N Clip Number	10		
Observations	4027		
Marginal R^2 / Conditional R^2	0.869 / 0.877		

4.2.2 Effects of RE type and group on accessibility score

Figure 6 summarizes the data for accessibility scores (ranging between -2 and 5) across all signers by RE Type.

As can be seen in Figure 6, for a referent that has a low accessibility score, the participants preferred to use nominals ($M_{\text{Accessibility}} = 0.29$), tracing ($M_{\text{Accessibility}} = 0.97$), and pronouns ($M_{\text{Accessibility}} = 1.33$). Note that while the majority of the tracing concentrated on the lower end of the accessibility scale (below 0), such constructions were also observed, albeit very low in number, toward the higher end, indicating that they could also be used for maintaining a previous

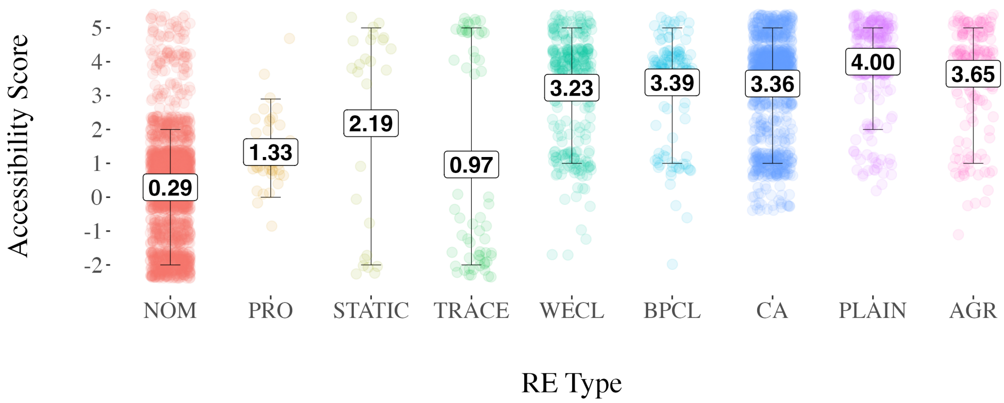


Figure 6. Mean Accessibility Score by RE Type. The brackets represent the four levels of RE Type entered into the statistical analysis. Note that pronouns do not belong to a group since they were omitted from the analysis due to their infrequent occurrence. The error bars represent the range between the 10th and 90th percentiles, whereas the text label represents the mean accessibility score

referent. Static specifiers ($M_{\text{Accessibility}} = 2.19$) were also very low in number but expressed referents with higher accessibility. WECL ($M_{\text{Accessibility}} = 3.23$), BPCL ($M_{\text{Accessibility}} = 3.39$), and CA ($M_{\text{Accessibility}} = 3.36$) were selected when they tracked highly accessible referents with similar mean accessibility scores. To compare, agreement ($M_{\text{Accessibility}} = 4.00$) and plain verbs ($M_{\text{Accessibility}} = 3.65$) had the highest accessibility scores, meaning that verbs almost only tracked referents that had the highest accessibility in the addressee's mind because of the lower inter-clausal distance to its previous mention, saliency in discourse, and zero or very few number of competitors.

Accessibility scores were analyzed with a mixed-effects regression model. Group (Native, Late) and RE Type (NOM, CL, CA, VERB) were entered as fixed effects, and participant and clip number were random effects. Multiple models were fit based on the number of predictors. Analyses included simpler models with only one predictor and progressively incorporated more complex models up to two-way interactions between Group and RE Type. Likelihood ratio tests were conducted to compare the fitted models. The inclusion of the Group improved the model's fit, $\chi^2(1) = 9.10, p = .003$. The model with a two-way interaction of Group and RE

Type as in (2) was the best-fitting model ($\chi^2(3) = 9.61$, $p = .022$), and its results will be reported below.

(2) $Model_2$ = lmer(Accessibility Group * RE Type + (1 | Participant) + (1 | Stimuli))[11]

The results of the model with RE Type as a predictor are presented in Table 4.

Table 4. Model results for Accessibility Score by Group and RE Type

	Accessibility Score		
Predictors	Estimates	CI	p
Intercept	2.58	2.38 – 2.78	<0.001
Group [Native]	0.16	0.08 – 0.24	<0.001
RE Type [NOM]	-2.29	-2.36 – -2.22	<0.001
RE Type [CL]	0.33	0.21 – 0.45	<0.001
RE Type [CA]	0.75	0.66 – 0.84	<0.001
Group [Native] * RE Type [NOM]	-0.11	-0.18 – -0.04	0.003
Group [Native] * RE Type [CL]	0.10	-0.01 – 0.21	0.076
Group [Native] * RE Type [CA]	0.01	-0.08 – 0.10	0.811
Random Effects			
σ^2	2.35		
τ_{00} Participant	0.03		
τ_{00} Clip Number	0.09		
ICC	0.05		
N Participant	28		
N Clip Number	10		
Observations	4027		
Marginal R^2 / Conditional R^2	0.498 / 0.521		

Due to the use of sum contrasts, the parameter estimate for one level of each categorical variable ("Verb" for RE Type and "Late" for Group) is not independently estimated and does not appear directly in the model output. Instead, it can be calculated based on the sum-to-zero constraint of the coding scheme. The results indicate that, among the grouped RE types, NOM significantly decreased the accessibility scores ($\beta = -2.29$, $p < .001$), indicating a decrease in accessibility. CL showed a modest increase in accessibility scores ($\beta = 0.33$, $p < .001$) compared to the grand mean, while CA was associated with a more substantial increase in accessibility ($\beta = 0.75$, $p < .001$).

[11] The string * indicates an interaction term.

To investigate whether there are significant differences between each RE Type level in the amount of accessibility score change, pairwise comparisons were calculated (Table 5). By conducting these pairwise comparisons for RE Type, it becomes possible to propose an accessibility hierarchy of forms, which will be further elaborated in Chapter 6.

Table 5. RE Type pairwise comparison results

Comparison	Estimates	SE	Z Ratio	p
CL - NOM	2.620	0.0797	32.866	<0.001
CL - CA	-0.417	0.0889	-4.691	<0.001
CL - VERB	-0.881	0.0995	-8.857	<0.001
NOM - CA	-3.037	0.0607	-50.004	<0.001
NOM - VERB	-3.501	0.0762	-45.950	<0.001
CA - VERB	-0.464	0.0840	-5.523	<0.001

The results in Table 5 indicate significant differences in accessibility scores among the different RE types, with all p-values less than 0.001. Pairwise comparisons show that CLs are associated with significantly higher accessibility scores than NOMs ($\beta = 2.620$), suggesting that CLs are more likely to be used when referents have higher accessibility compared to NOMs. Conversely, CAs and verbs are associated with even higher accessibility scores than CLs, as indicated by negative estimates of -0.417 (CL vs. CA) and -0.881 (CL vs. verbs), implying that CAs and verbs are more frequently used with referents of higher accessibility than CLs are. Further comparisons reveal that NOMs are associated with substantially lower accessibility scores than CAs and verbs, with estimates of -3.037 and -3.501, respectively. This reinforces that NOMs are more commonly used with referents of lower accessibility compared to CAs and verbs. The comparison between CAs and verbs ($\beta = -0.464$) indicates that verbs are associated with slightly higher accessibility scores than CAs, though the difference is smaller compared to other contrasts. Overall, both verbs and CAs are used with referents of higher accessibility, with verbs being slightly more so.

Finally, Figure 7 presents accessibility scores for Group by RE Type and Discourse.

Figure 7 demonstrates that native signers (blue circular shape) overall referred to entities that had more accessibility for most RE types than late signers (green triangular shape). One striking difference in the introduced context is that late signers diverged from the native signers when they used tools such as BPCL ($M_{\text{Accessibility}} = 1.00$), CA ($M_{\text{Accessibility}} = 0.50$), and other verbal constructions like PLAIN ($M_{\text{Accessibility}} = 1.00$) to introduce referents with relatively higher accessibility for referent introduction purposes. Native signers either did not use such tools for referent introduction (e.g., CA and PLAIN) or used similar tools for less accessible referents (BPCL: $M_{\text{Accessibility}} = -1.50$; AGR: $M_{\text{Accessibility}} = -1.00$). The average accessibility score for the maintained and re-introduced referents among both groups was similar, with native signers tracking referents with slightly higher accessibility. Moreover, in the maintenance and re-introduction contexts, both native and late signers demonstrated similar usage of RE types, with comparable mean accessibility scores, indicating a convergence in the tracking of highly accessible referents. To illustrate, in the maintenance discourse context, both groups tracked referents who were assigned high mean accessibility scores ranging from approximately 4.0 to 4.75. In the re-introduction context, this pattern persisted, with both groups tracking referents with relatively higher mean accessibility scores, albeit with native signers occasionally exhibiting slightly higher values for the referents. For example, native signers used AGR to re-introduce slightly more accessible referents ($M_{\text{Accessibility}} = 1.42$) compared to late signers ($M_{\text{Accessibility}} = 1.04$). Similarly, with PLAIN, native signers tracked more accessible referents ($M_{\text{Accessibility}} = 1.50$) for re-introduction, whereas late signers used the same tool to track slightly less accessible referents ($M_{\text{Accessibility}} = 1.09$).

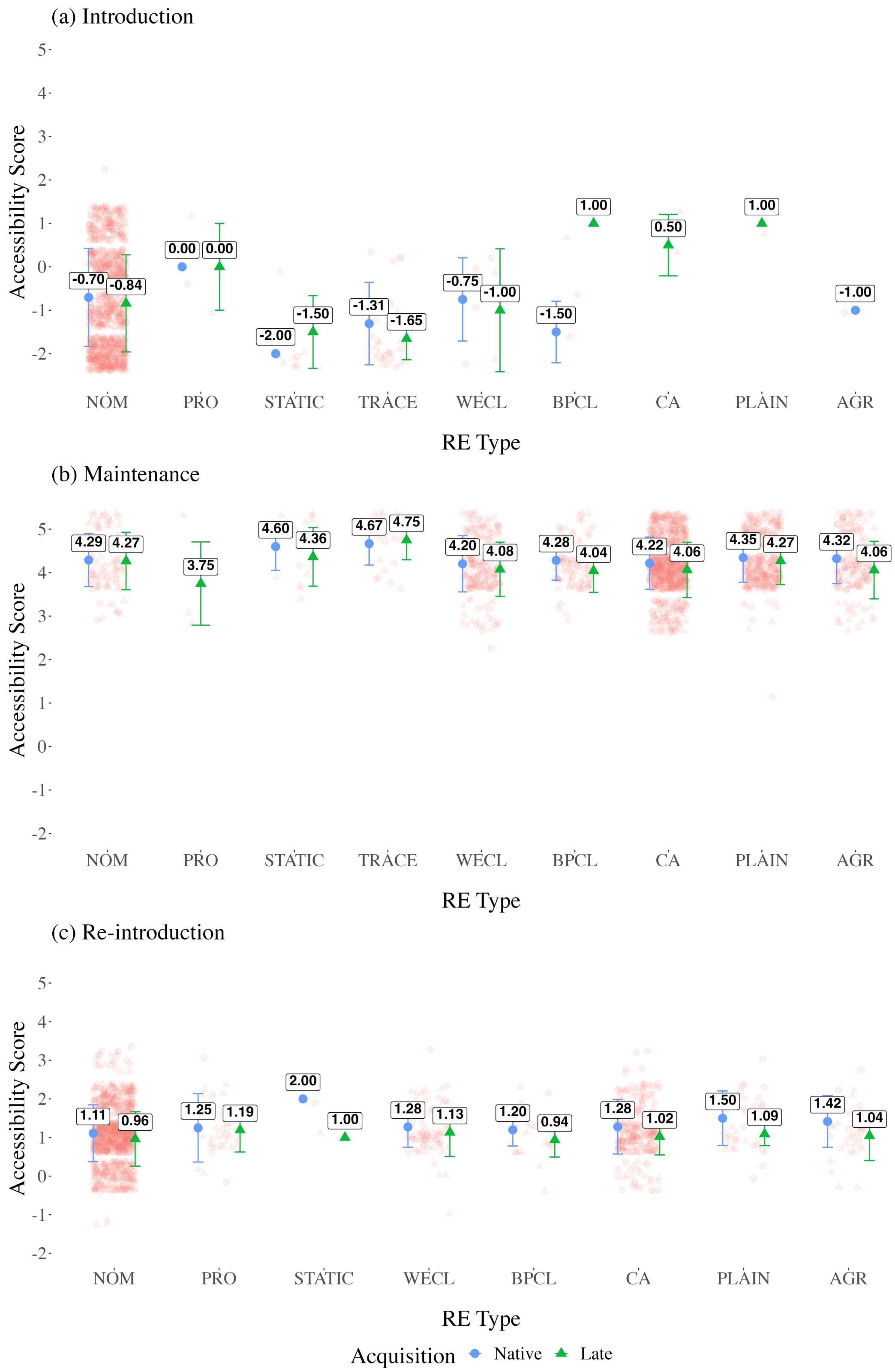


Figure 7. Accessibility Score by RE Type, Group and Discourse: Introduction (a), Maintenance (b) and Re-introduction (c). The brackets represent the four levels of RE Type entered into the statistical analysis. Shapes represent the mean accessibility score

As per the group effect, the model results in Table 4 show that being a native signer increased the accessibility scores significantly, with native signers having significantly higher accessibility than late signers, $\beta = 0.16$, 95% CI: 0.08 – 0.24, $p < .001$. As for the interaction effect between Group (Native) and RE Type (NOM), it is possible to see that the negative effect of NOM on accessibility was more pronounced for native signers ($\beta = -0.11$, CI: -0.18 – -0.04, $p = .03$). A negative beta coefficient in the interaction term indicates that the decrease in accessibility associated with NOM is greater among native signers compared to late signers. In other words, native signers used NOM as an accessibility-decreasing tool more frequently than late signers. Additionally, the second coefficient in the interaction terms, which is marginally non-significant, indicates that CLs are less of an accessibility-decreasing tool for native signers compared to late signers ($\beta = 0.10$, CI: -0.01 – 0.21, $p = .076$). In other words, the decreasing impact of CL on accessibility is mitigated among natives. This suggests that native signers have a higher accessibility threshold than late signers for employing CLs, requiring referents to be more accessible before using this type of RE. Since CL is used mainly as an accessibility-increasing tool, necessitating higher levels of accessibility scores to use this tool means better compliance to principles of discourse cohesion by natives. Other interaction terms were not statistically significant, which indicates that the effects of most RE types on accessibility did not differ markedly between native and non-native signers.

As an interim summary, the analyses presented in Sections 4.2.1 and 4.2.2 highlight the effects of Discourse, RE Type, and Group on the accessibility scores of referents in TİD narratives. The findings demonstrate that discourse context is a strong predictor of accessibility scores. Specifically, introduced referents have the lowest accessibility ($M = -0.80$), re-introduced referents have moderate accessibility ($M = 1.08$), and maintained referents possess the highest accessibility ($M = 4.2$). This pattern aligns with theories of linguistic efficiency and referential accessibility, suggesting that signers use more explicit forms for less accessible referents and more reduced forms for highly accessible ones.

Regarding RE Type, the results indicate that nominals are predominantly used to introduce less accessible referents, while classifiers and constructed action are employed for referents with moderate accessibility. Verbal constructions, including agreement and plain verbs, are primarily used to track highly accessible referents during maintenance. The statistical models showed that RE Type significantly predicts accessibility scores, although not as strongly as discourse context. Furthermore, the pairwise comparisons revealed that all four levels significantly differed from each other in terms of referential accessibility. Finally, the group analysis reveals that native signers tend to track more accessible referents than late signers, indicating subtle differences in discourse cohesion due to the age of acquisition. Native signers use more implicit forms, such as constructed action and verbs, especially for reference maintenance and re-introduction. In contrast, late signers rely more on explicit forms like nominals, even when referring to highly accessible referents. Statistical interactions between Group and RE Type suggest that late signers may be slightly over-explicit in their reference tracking, possibly due to differences in pragmatic sensitivity.

CHAPTER 5

PHONETIC ANALYSIS

5.1 Landmark extraction and data preprocessing

For the landmark extraction phase, I used the Python (3.12.4) implementation of MediaPipe (0.10.14) (Lugaresi et al., 2019), OpenCV (4.10.2), and NumPy (2.0.0) to detect and track body landmarks. This extraction involves identifying spatiotemporal key points such as wrists, elbows, and shoulders or hands for both the left and right sides of the body. All signers in the videos were right-handed (i.e., their dominant hand was their right hand). The 2D videos were inputted into the MediaPipe package frame-by-frame. The input videos were recorded with different rates of frames per second (fps). As a result, during video extraction, all clips were transformed to have 50 fps with `ffmpeg` (7.0.1). The input videos were the segmented FARE and KEDI nouns in the RE analysis, and they were the filtered nine events in the event analysis. The key points from the signer's body were then tracked, yielding a skeleton representation with numerical coordinates in a three-dimensional space (x, y, z) per frame (see Figure 8 for details).

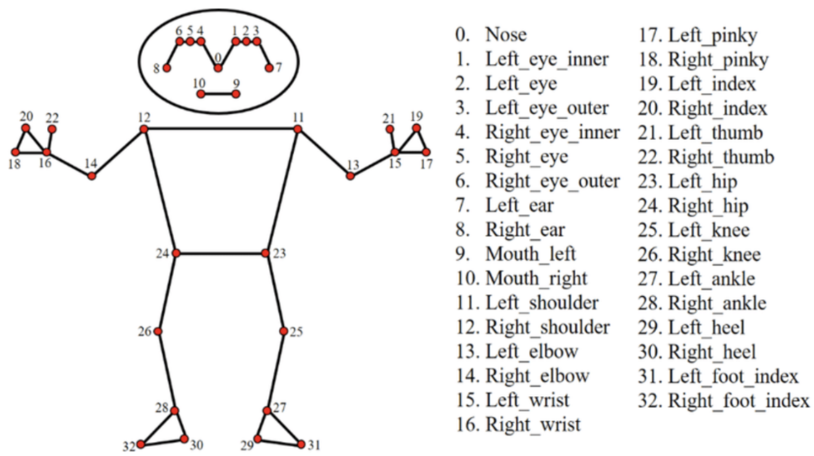


Figure 8. The 33 landmarks used for body pose estimation in the MediaPipe framework (Lugaresi et al., 2019).

MediaPipe's pose estimation model originally outputs 33 triplets (x, y, z), each of which represents an estimated body landmark, but here I only focus on 10 of them: left thumb, right thumb, left pinky, right pinky, left wrist, right wrist, left elbow, right elbow, left shoulder and right shoulder. The illustration of these extracted key points is illustrated in Figure 9. Before using these values to calculate phonetic measures, I normalized the extracted triplet data.

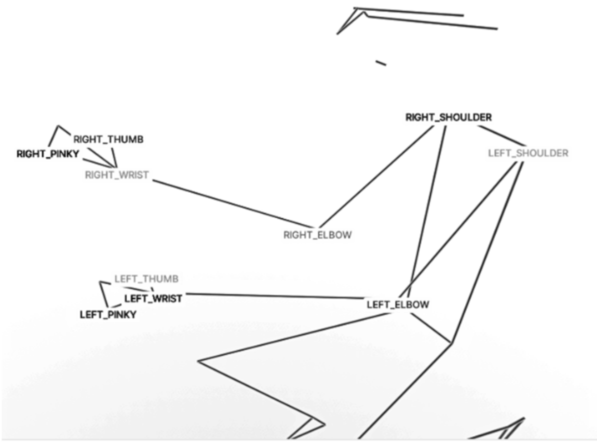


Figure 9. The ten extracted body landmarks from MediaPipe were mapped to 3D space

During preprocessing, the raw RE and event data underwent several transformations to ensure reliability and consistency across the dataset. I adopted the following preprocessing steps from previous research (Börstell, 2023; Flaherty, Sato, & Kirby, 2023; Östling, Börstell, & Courtaux, 2018; Stamp, Cohn, Hel-Or, & Sandler, 2024; Stamp, Dachkovsky, Hel-Or, Cohe, & Sandler, 2024). First, a median filter calculated through the equation in (1) was applied to smooth out the noise in the landmark tracking data, minimizing the effect of sudden, aberrant changes in position due to the tracking software or the signer's rapid movements:

(1) $Y_i = \text{median}(X_{i-n}, \ldots, X_{i+n})$

where Y_i represents the filtered value, X_i is the original data point, and n is the number of points considered in the filter window.

The data were normalized to eliminate differences due to individual signer body dimensions, using the mean upper arm length as a scaling factor. I took the average Euclidean distance, as shown in (2), between the left and right shoulder and elbow joints to calculate the upper arm length for each participant. Following Flaherty et al. (2023), the data was scaled by using a specific scaling factor calculated for each participant. This scaling factor was determined by dividing each participant's upper arm length by the average upper arm length across all participants, allowing adjustments based on individual differences in arm length.

(2) $$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

where D represents the distance and the coordinates of the compared joints.

Following the normalization, I computed the following phonetic measures:

1) Duration: This is the interval in seconds (sec) from the first frame (onset) to the last frame (offset) of the sequence (noun referent or event).

2) Accumulated Hand Distance: This is the normalized accumulated Euclidean distance calculated through the equation in (2) between the coordinates of the hand joint in consecutive frames. Since MediaPipe does not have a specific landmark for "hand," I had to calculate an estimation of this joint's coordinates based on the extracted thumb and pinky landmarks. I took the average of thumb and pinky coordinates for each axis for both sides (right and left). Equation in (3) shows this calculation.

(3) $$(X_{\text{HAND}}, Y_{\text{HAND}}, Z_{\text{HAND}}) = \left(\frac{X_{\text{THUMB}} + X_{\text{PINKY}}}{2}, \frac{Y_{\text{THUMB}} + Y_{\text{PINKY}}}{2}, \frac{Z_{\text{THUMB}} + Z_{\text{PINKY}}}{2}\right)$$

To calculate Accumulated Hand Distance, I used the Euclidean distance formula (2) between two consecutive right-hand positions in a 3-dimensional space. This calculates the movement rate by measuring the change in position between two-time points or frames, t and $t - 1$, along the X, Y, and Z axes (4). Then, per each RE or event, I sum the differences to obtain Accumulated Hand Distance, representing the total amount of movement in one RE or event. The calculation is illustrated in Figure 10.

(4)

$$D_{\text{HAND}} = \sqrt{\begin{aligned}(X_{\text{HAND}}(t) - X_{\text{HAND}}(t-1))^2 + \\ (Y_{\text{HAND}}(t) - Y_{\text{HAND}}(t-1))^2 + \\ (Z_{\text{HAND}}(t) - Z_{\text{HAND}}(t-1))^2\end{aligned}}$$

where D_{HAND} is the hand distance between consecutive frames at time t and $t-1$.

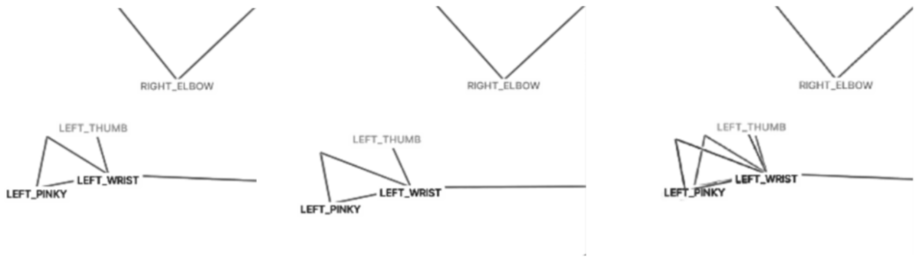


Figure 10. Pictures 1 and 2: Consecutive positions (in frames 111 and 112) of the hand landmarks. Picture 3: Both frames overlaid on top of each other, indicating the amount of distance change

3) Duration-Normalized Accumulated Hand Distance: The summed hand distances per referent or event were averaged and normalized across the total duration of each video to provide a clear, continuous metric of movement relative to the signer's body. The normalization of distance over time is provided in the formula in (5):

(5) $\text{Normalized Distance} = \frac{\sum_{i=1}^{N} D_i}{T_{\text{max}}}$

where D_i is the individual distance at each frame, N is the number of frames, and T_{max} is the maximum time duration (in seconds) of the video.

4) Duration-Normalized Sign Space Use: This is the Euclidean distance shown in (2) after normalization between the wrist joint and a reference point on the body (the joint tracked between the shoulders). Since MediaPipe does not have a landmark for mid-shoulder, I estimated the position of this point by averaging left and right shoulders for each axis, as shown in (6).

(6)

$$X_{\text{M\_SHOULDER}} = \frac{X_{\text{L\_SHOULDER}} + X_{\text{R\_SHOULDER}}}{2},$$
$$Y_{\text{M\_SHOULDER}} = \frac{Y_{\text{L\_SHOULDER}} + Y_{\text{R\_SHOULDER}}}{2},$$
$$Z_{\text{M\_SHOULDER}} = \frac{Z_{\text{L\_SHOULDER}} + Z_{\text{R\_SHOULDER}}}{2},$$

where M is Mid, L is for Left, and R is for Right.

After obtaining the values for the point between the two shoulders, I calculated the Euclidean distance between the wrists (left and right separately) to the reference point anchored on the body, which is the mid-shoulder point for each frame (7).

(7)

$$D_{\text{SHOULDER\_WRIST}} = \sqrt{(X_{\text{WRIST}} - X_{\text{MID\_SHOULDER}})^2 + (Y_{\text{WRIST}} - Y_{\text{MID\_SHOULDER}})^2 + (Z_{\text{WRIST}} - Z_{\text{MID\_SHOULDER}})^2}$$

I then, calculated the change in distance (absolute difference) between two consecutive measurements and for the shoulder-to-wrist distance as in (8) and then summed the distances per each RE and event. This represents the magnitude of change in the use of signing space throughout the utterance.

(8) $\Delta D_{\text{SHOULDER\_WRIST}} = |D_{\text{SHOULDER\_WRIST}}(t) - D_{\text{SHOULDER\_WRIST}}(t-1)|$

Finally, I normalized the values by dividing the distances by the duration in (5). Overall, this measurement is meant to be taken as a proxy for the use of the signing space. A more expansive use of the signing space per RE or event will correspond to a larger change in distance between shoulder and wrists, whereas a narrower use of the space will signal a more minor change. This calculation is illustrated in Figure 11.

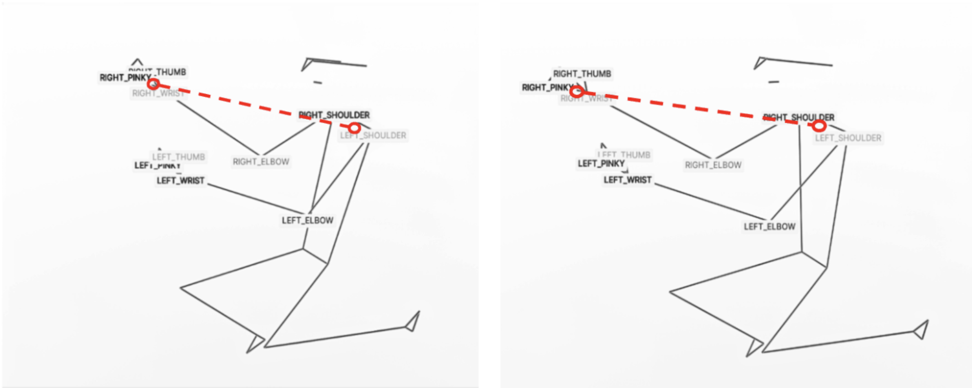


Figure 11. Calculation of the sign space use. The red line shows an approximation of the distance between consecutive positions in frames 111 (left) and 112 (right) of the right wrist and mid-shoulder

5.2 RE analysis

5.2.1 Statistical data analysis

Following the data transformation and calculation of the three phonetic measures, the .csv file outputted by the MediaPipe package was transferred to R for statistical analyses, and the same preprocessing steps (median filtering and body normalization) were applied. Key metrics were calculated, such as average accumulated hand distance and sign space use (left and right sides).

Initial normality checks revealed deviations from the Gaussian distribution for each of the three phonetic measures: duration ($W = .74, p < .001, D = .53, p < .001$), accumulated hand distance ($W = .87, p < .001, D = .59, p < .001$), sign space use ($W = .90, p < .001, D = .52, p < .001$). As a result, all of the three phonetic measures were log-transformed (Feng et al., 2014). Furthermore, outlier data points beyond ± 3 times the interquartile range were removed from the log-transformed data, corresponding to 2% of the data. After these steps, the data showed better normality (duration: $W = .99, p < .01, D = .54, p < .001$; accumulated hand distance: $W = .99, p < .001, D = .24, p < .001$; sign space use: $W = .99, p < .001, D = .76, p < .001$). Each log-transformed phonetic measure was analyzed with mixed effects models using the `lmer` function of the `lme4` package (Bates et al., 2024).

Mixed effects models were built to analyze the phonetic measures. The outcome variables were the phonetic measures (Duration, Duration-Normalized Accumulated Hand Distance, and Duration-Normalized Sign Space Use). One set of models had Discourse (Introduction, Maintenance, and Re-introduction) and RE (FARE and KEDI) as fixed effects, and Participant and Stimuli Number as random effects. The other set of models had Accessibility Score (ordinal) and RE (FARE and KEDI) as fixed effects to examine this effect separately. The same random effects were entered into the model.

5.2.2 Results

I analyzed 23564 frames pertaining to the extracted nominal REs of FARE and KEDI with MediaPipe's pose landmark estimation. 47% of the frames (11108 in number) were occurrences of FARE (Introduction: 5431 frames, Maintenance: 468 frames, Re-introduction: 5209 frames), while the remaining 53% of the frames (12456 in number) were KEDI (Introduction: 6182 frames, Maintenance: 781 frames, Re-introduction: 5493). As described in Section 5.1, all data points went through certain steps of normalization and preprocessing prior to analysis. The normalized phonetic values were then grouped by participant, stimuli number, RE type, and RE order. This also corresponded to a total of 1055 unique data points for statistical analysis. As a result of the outlier removal described above, the results report on a total of 1033 data points (FARE: 468, KEDI: 565). 404 of them were introduced REs, 69 were maintained, and the remaining 560 were re-introduced.

Out of these 1033 nominal REs (KEDI and FARE), 444 instances tracked referents with an accessibility score of 1. This was followed by a score of 0, with 248 RE occurrences, and a score of 2, with 134 REs. Lower accessibility scores, such as -1 and -2, were less frequent, with 103 and 23 instances. Since signers generally do not employ nominals to maintain a referent, the higher end of the accessibility scale is less densely populated. Accessibility scores of 3 and 4 were observed 22 and 32 times in this data, respectively, and a score of 5 was observed only in 27 instances.

In the analyses, the left and right of the body were analyzed separately. All the phonetic measures (hand distance and space use) I report henceforth are the log-transformed averages of both sides.

Figure 12 displays the results for RE duration (given in seconds) by discourse. In the figure, the red line indicates the mean duration across the three levels of discourse. The dashed black line indicates linear regression.

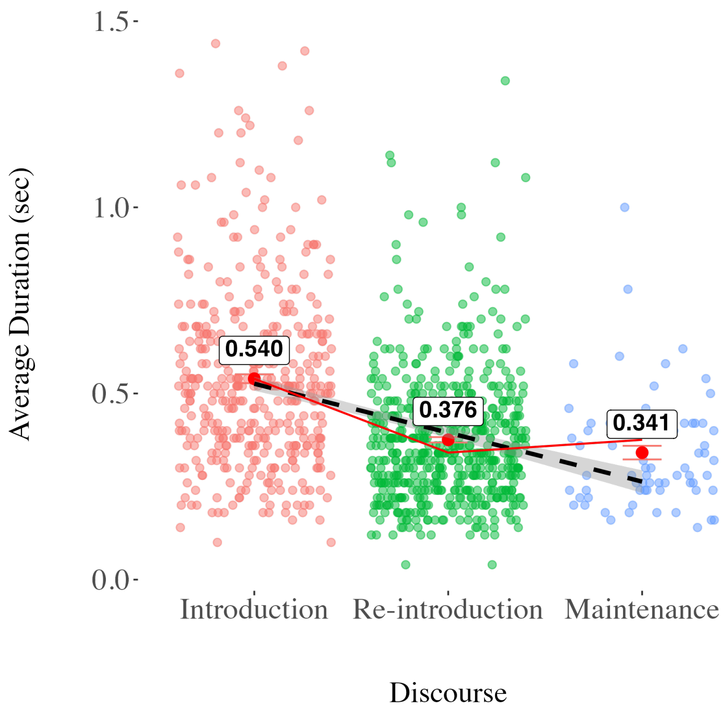


Figure 12. Average RE Duration (sec) by Discourse. Each dot is a single observation (RE). The red line and points represent the mean observed value for each bin. The dashed black line is a regression line fitted across the combined data from all discourse contexts

In Figure 12, it can be seen that the mean duration of REs was the highest for introduction ($M = .54$ seconds, $SE = .012$), compared to re-introduction ($M = .38$ seconds, $SE = .007$) and maintenance ($M = .34$ seconds, $SE = .018$). This means that signers spent more time articulating REs when introducing new referents into the discourse.

Figure 13 presents duration-normalized accumulated hand distance by discourse.

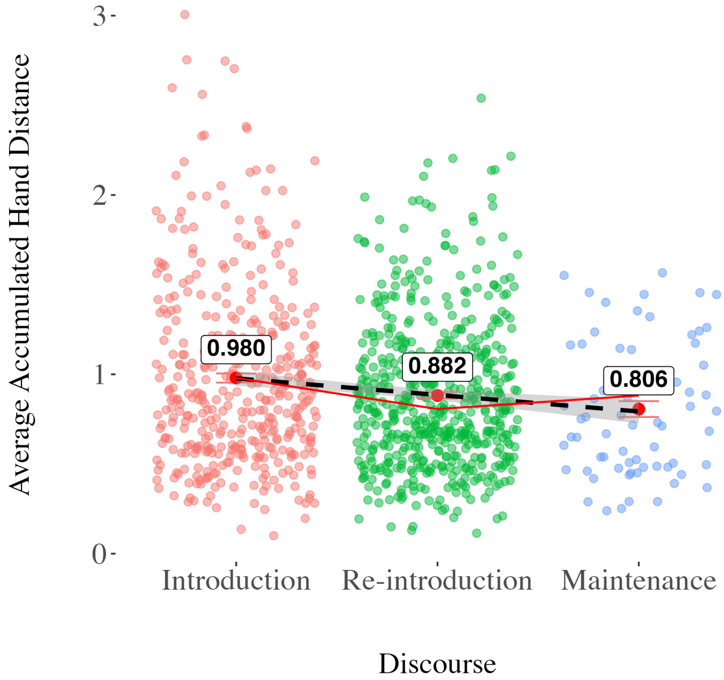


Figure 13. Average Duration-Normalized Hand Distance Use by Discourse. Each dot is a single observation (RE). The red line and points represent the mean observed value for each bin. The dashed black line is a regression line fitted across the combined data from all discourse contexts

The accumulated hand distance, normalized by duration, was the greatest in the introduction context ($M = .98$, $SE = .025$), followed by re-introduction ($M = .88$, $SE = .017$) and maintenance ($M = .81$, $SE = .044$).

Figure 14 shows the mean normalized sign space use (the tracked distance between wrists and middle shoulder point) by discourse.

The size of the signing space was also the largest in referent introduction ($M = .24$, $SE = .005$), followed by re-introduction ($M = .23$, $SE = 0.004$) and maintenance ($M = .22$, $SE = .010$), meaning that signers overall used a narrower sign space to track more accessible referents. In contrast, during the introduction of referents into the context, signers employed longer durations, larger hand movements or distance, and a wider use of signing space.

Figure 15 presents the duration analyzed across different levels of referent accessibility.

RE duration exhibited variability across the accessibility score spectrum. Generally, referents with lower accessibility scores, such as -2 ($M = .41$ sec,

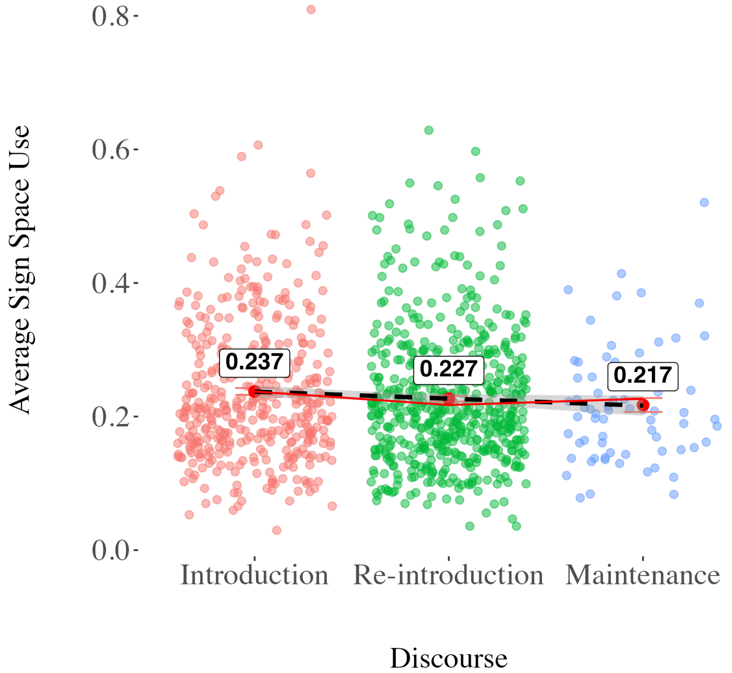


Figure 14. Average Duration-Normalized Sign Space Use by Discourse. Each dot is a single observation (RE). The red line and points represent the mean observed value for each bin. The dashed black line is a regression line fitted across the combined data from all discourse contexts

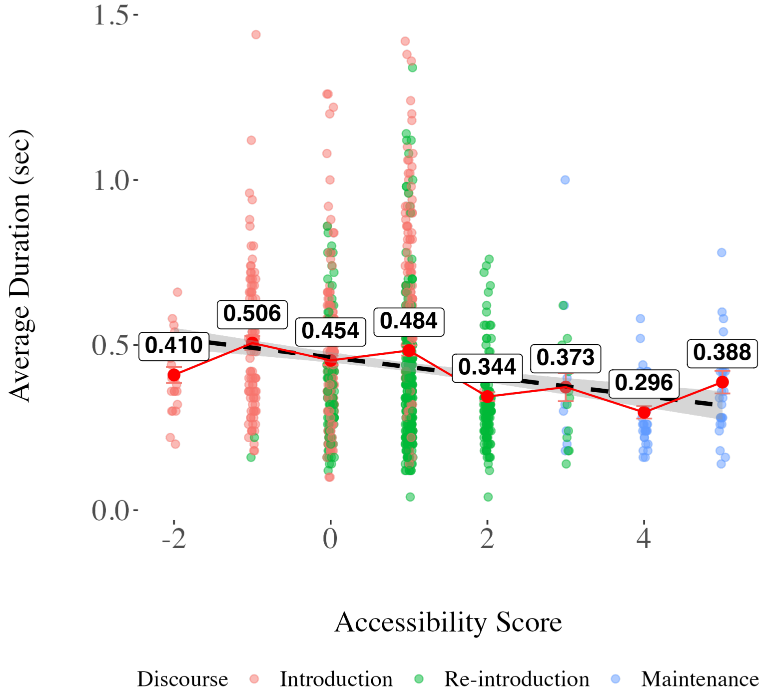


Figure 15. Average RE Duration by Accessibility Score. Each dot is a single observation (RE). The red line and points represent the mean observed value for each bin. The dashed black line is a regression line fitted across the combined data from all accessibility bins

SE = .024 sec), −1 (M = .51 sec, SE = .021 sec), and 0 (M = .45 sec, SE = .14 sec) were associated with longer durations. Conversely, referents with higher accessibility scores, namely 3 (M = .37 sec, SE = .043 sec) and 4 (M = .30 sec, SE = .019 sec), had shorter durations, except for the referents with the accessibility score of 5 (M = .39 sec, SE = .034 sec) where the duration was longer compared to other accessible levels.

Figure 16 displays accumulated hand distance as a function of accessibility scores.

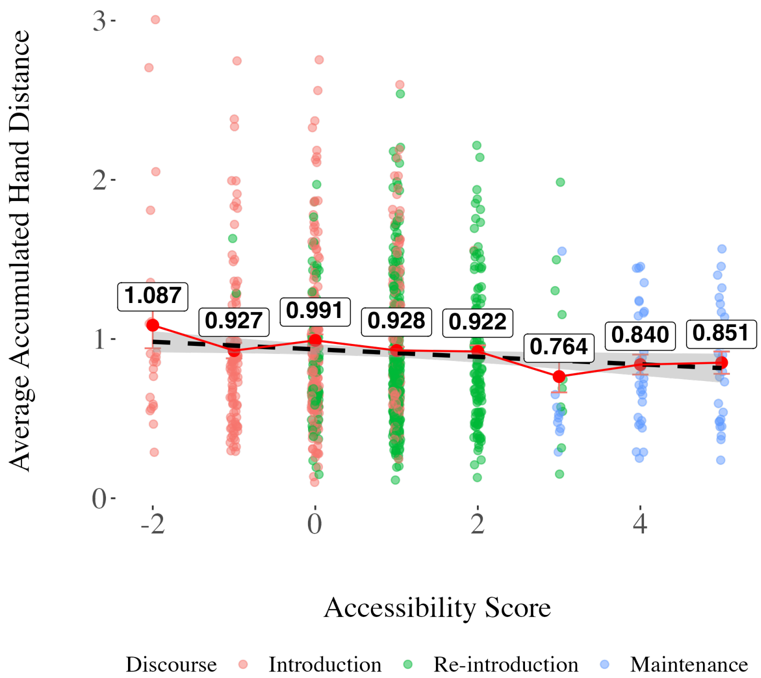


Figure 16. Average Duration-Normalized Hand Distance by Accessibility Score. Each dot is a single observation (RE). The red line and points represent the mean observed value for each bin. The dashed black line is a regression line fitted across the combined data from all accessibility bins

Hand distance generally decreased as referent accessibility increased. Referents with an accessibility score of −2 had the highest mean hand distance (M = 1.09, SE = .145), followed by other low accessible scores such as −1, 0, 1, and 2 (M ranged between .92 and .99). Then, there is an identifiable reduction in hand distance with an accessibility level of 3 (M = .76, SE = .099). A similar decrease was also observed for other high accessibility scores, such as 4 (M = .84, SE = .063) and 5 (M = .85, SE = .069), albeit to a lesser extent.

Lastly, Figure 17 represents sign space use as a function of accessibility scores.

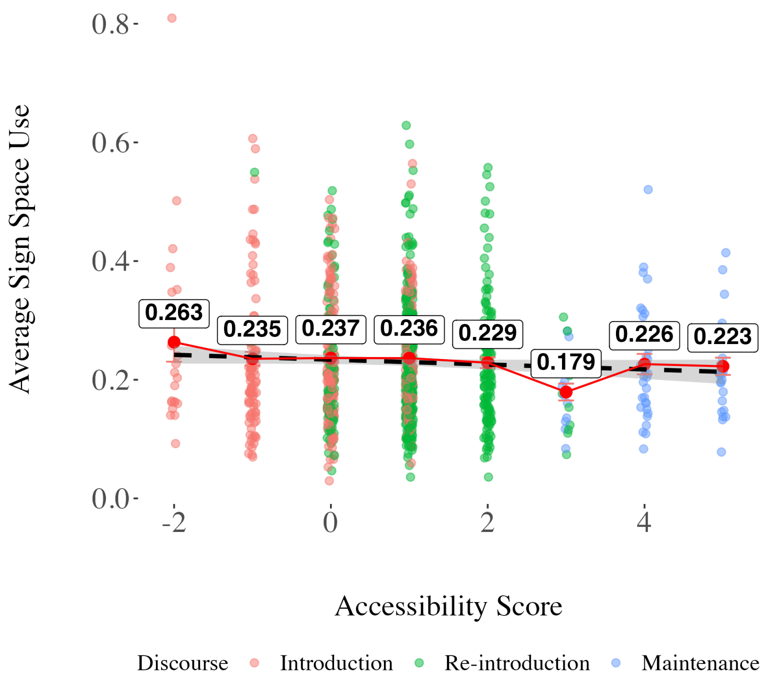


Figure 17. Average Duration-Normalized Sign Space Use by Accessibility Score. Each dot is a single observation (RE). The red line and points represent the mean observed value for each bin. The dashed black line is a regression line fitted across the combined data from all accessibility bins

Sign space use showed a slight decreasing trend with increasing accessibility. One noticeable difference in sign space occurred when switching from the accessibility score of -2 to -1, in which the RE is signed in a narrower space on average with increasing accessibility (-2: $M = .26$, $SE = .033$; -1: $M = .24$, $SE = .012$). From the accessibility level -1 to 2, the sign space use remained relatively constant (M ranging from .23 to .24). This is followed by another noticeable reduction in space use for the accessibility score of 3 ($M = .18$, $SE = .014$), and increases again for the remaining two levels of accessibility (4: $M = .23$, $SE = .018$; 5: $M = .22$, $SE = .015$). Overall, the phonetic measures analyzed across both discourse contexts and accessibility levels reveal a linear pattern: less accessible referents and those introduced into the discourse are associated with increased phonetic effort, as evidenced by longer durations, greater hand movements, and larger sign space use. In contrast, more accessible referents and those maintained in the discourse are expressed with reduced phonetic effort.

Mixed-effects models were used to analyze the phonetic measures for both Discourse and Accessibility. To choose the best possible model (among one and two predictor models), likelihood ratio tests were conducted to compare multiple models for both sets of models. For the duration, the two-predictor model that also included RE was significantly better than the simpler model both in the discourse set ($\chi^2(1) = 7.71, p = .005$) and in the accessibility set, $\chi^2(1) = 3.90, p = .048$. Furthermore, the model that had an interaction term between Discourse and RE had higher likelihood than the model that lacked the interaction for the discourse set ($\chi^2(2) = 10.80, p = .005$) but not for the accessibility set, $\chi^2(1) = .12, p = .729$. As a result, the two models that I will report below for duration are listed in (9) and (10).

For accumulated hand distance, one predictor model that included the main effect of Discourse (the first set) and Accessibility (the second set) was not a worse fit than a two predictor model, $\chi^2(1) = .03, p = .869$; $\chi^2(1) = .17, p = .681$, respectively. The formulas of the selected models that analyze hand distance are given in (11) and (12).

Finally, for the sign space analysis, the two predictor models were a better fit than one predictor models ($\chi^2(1) = 5.16, p = .023$ for the discourse set; $\chi^2(1) = 6.30, p = .012$ for the accessibility set). The interaction models had a lower likelihood ($\chi^2(2) = 3.79, p = .15$) for the discourse set but not for the accessibility set, $\chi^2(1) = 4.26, p = .039$. Following this, the chosen sign space models were formulated as in (13) and (14).

(9) $\text{Model}_{\text{Duration-1}}$ = lmer(log(Duration)   Discourse * RE + (1 | Participant) + (1 | Stimuli))

(10) $\text{Model}_{\text{Duration-2}}$ = lmer(log(Duration)   Accessibility + RE + (1 | Participant) + (1 | Stimuli))

(11) ($\text{Model}_{\text{Hand.Distance-1}}$ = lmer(log(Distance)   Discourse + (1 | Participant) + (1 | Stimuli))

(12) $\text{Model}_{\text{Hand.Distance-2}}$ = lmer(log(Distance)   Accessibility + (1 | Participant) + (1 | Stimuli))

(13) $\text{Model}_{\text{Sign.Space.Use-1}}$ = lmer(log(Distance)   Discourse + RE + (1 | Participant) + (1 | Stimuli))

(14) $\text{Model}_{\text{Sign.Space.Use-2}}$ = lmer(log(Distance)   Accessibility * RE + (1 | Participant) + (1 | Stimuli))

The results of the mixed-effects models of the log-transformed phonetic measures that include Discourse and/or RE as predictors are presented in Table 6. Since I use sum contrasts too in these models, the parameter estimate for one level of each categorical variable (Re-introduction for Discourse and KEDI for RE) is not included in the model results.

Table 6. Model results for RE-level analysis by Discourse and RE

	Duration (log)			Duration-Normalized Accumulated Hand Distance (log)			Duration-Normalized Sign Space Use (log)		
Predictors	Estimate	CI	p	Estimate	CI	p	Estimate	CI	p
Discourse [Introduction]	0.28	0.23 – 0.33	<0.001	0.09	0.04 – 0.14	0.001	0.06	0.01 – 0.10	0.010
Discourse [Maintainance]	-0.18	-0.25 – -0.11	<0.001	-0.06	-0.14 – 0.01	0.096	-0.03	-0.09 – 0.04	0.407
RE [FARE]	0.05	0.01 – 0.09	0.019				-0.03	-0.06 – -0.01	0.017
Discourse [Introduction] × RE [FARE]	-0.07	-0.12 – -0.03	0.002						
Discourse [Maintainance] × RE [FARE]	0.05	-0.02 – 0.12	0.187						
Random Effects									
σ^2	0.16			0.21			0.16		
τ_{00} Participant	0.04			0.04			0.05		
τ_{00} Narrative	0.01			0.01			0.00		
ICC	0.23			0.20			0.25		
N Participant	28			28			28		
N Narrative	10			10			10		
Observations	1033			1033			1033		
Marginal R^2 / Conditional R^2	0.169 / 0.359			0.012 / 0.214			0.013 / 0.256		

The model results show that the discourse context significantly affects the duration of REs and support the empirical observations I described in the previous paragraphs. Specifically, referent introduction significantly increases the duration of REs ($\beta = 0.29$, CI: 0.24 – 0.33, $p < 0.001$). Conversely, referent maintenance significantly decreases the duration ($\beta = -0.18$, CI: -0.25 – -0.11, $p < 0.001$). The effect of RE [FARE] was also significant ($\beta = 0.05$, CI: -0.01 – 0.09, $p = 0.016$), indicating a trend where REs referring to FARE might have slightly longer durations compared to KEDI. This result stems from the fact that these two signs have different phonological compositions (e.g., movement, handshape, and orientation). When we look at the interaction term, a significant interaction was found between Discourse [Introduction] and RE [FARE] ($\beta = -0.06$, CI: -0.11 – -0.02, $p = 0.008$). This suggests that the increase in duration for introduced referents is less pronounced for FARE than KEDI. In other words, when introducing the referent FARE, signers do not lengthen the duration as much as they do for KEDI, which could be again attributed to sign-internal phonological composition. Nevertheless, the interaction between Discourse [Maintenance] and RE [FARE] was not significant ($\beta = 0.04$, CI: -0.03 – 0.11, $p = .282$), indicating that during maintenance, the duration of REs does not differ significantly between FARE and KEDI.

As for hand distance, the results of the single predictor model indicated a positive effect of Discourse [Introduction] ($\beta = 0.04$, CI: 0.04 – 0.14, $p < .001$). Overall, introducing a nominal RE increased the amount of hand movement. Maintaining it had a decreasing effect; however, it did not reach statistical significance ($\beta = -0.07$, CI: -0.15 – 0.01, $p = .080$).

Additionally, the two predictor model for sign space use revealed that introduced nominal REs significantly increased the distance between wrist and mid-shoulder point ($\beta = 0.06$, CI: 0.01 – 0.10, $p < .009$) while the decreasing effect of maintenance was again insignificant ($\beta = -0.03$, CI: -0.10 – 0.03, $p = .332$). Finally, the second predictor, RE [FARE], had a significant effect. The use of FARE had a negative or reduction effect on the signing space ($\beta = -0.03$, CI: -0.06 – -0.00, $p < .023$).

The mixed-effects regression models presented in Table 7 further examine the relationship between referent accessibility and the phonetic measures of REs. Like the previous analysis, each model used one of the log-transformed phonetic measures as the dependent variable: duration, duration-normalized accumulated hand distance, and duration-normalized sign space use.

Table 7. Model results for RE-level analysis by Accessibility and RE

	Duration (log)			Duration-Normalized Accumulated Hand Distance (log)			Duration-Normalized Sign Space Use (log)		
Predictors	Estimate	CI	p	Estimate	CI	p	Estimate	CI	p
Accessibility [linear]	-0.35	-0.51 – -0.20	<0.001	-0.21	-0.38 – -0.05	0.013	-0.16	-0.30 – -0.02	0.030
Accessibility [quadratic]	0.01	-0.14 – 0.16	0.900	0.17	0.01 – 0.32	0.037	0.17	0.03 – 0.30	0.014
Accessibility [cubic]	0.21	0.06 – 0.36	0.005	0.07	-0.09 – 0.22	0.398	0.06	-0.08 – 0.19	0.403
Accessibility [4th degree]	-0.02	-0.14 – 0.10	0.700	0.04	-0.09 – 0.17	0.540	0.03	-0.08 – 0.14	0.560
Accessibility [5th degree]	0.06	-0.06 – 0.18	0.311	-0.11	-0.24 – 0.02	0.084	-0.10	-0.21 – 0.01	0.064
Accessibility [6th degree]	0.01	-0.11 – 0.13	0.887	-0.09	-0.22 – 0.04	0.168	-0.09	-0.20 – 0.02	0.110
Accessibility [7th degree]	0.15	0.06 – 0.24	0.001	-0.04	-0.14 – 0.05	0.368	-0.02	-0.10 – 0.06	0.563
RE1	0.03	0.01 – 0.06	0.022				-0.03	-0.06 – -0.01	0.013
Random Effects									
σ^2	0.19			0.21			0.16		
τ_{00} Participant	0.03			0.05			0.05		
τ_{00} Narrative	0.01			0.01			0.00		
ICC	0.17			0.21			0.25		
N Participant	28			28			28		
N Narrative	10			10			10		
Observations	1033			1033			1033		
Marginal R^2 / Conditional R^2	0.069 / 0.231			0.012 / 0.221			0.016 / 0.261		

For the duration of REs, the model showed a significant linear negative effect of accessibility scores ($\beta = -0.35$, CI: -0.51 – -0.20, $p < .001$). This indicates that as referent accessibility increases by one unit, the log-transformed duration of REs linearly decreases by 0.07 units on average. In practical terms, more accessible referents are produced with shorter durations. The main effect of RE was also significant in the model. FARE increased the duration, $\beta = -0.07$, CI: -0.09 – -0.05, $p < .001$. Regarding the duration-normalized accumulated hand distance, accessibility again had a linear significant negative effect ($\beta = -0.21$, CI: -0.38 – -0.05, $p = .013$). In other words, TİD signers used less hand movement when signing more accessible nominals. Finally, for the duration-normalized sign space use, accessibility also exhibited a significant negative effect ($\beta = -0.16$, CI: -0.30 – -0.02, $p = .030$). This result indicates that higher accessibility scores are associated with a slight reduction in the signing space. TİD signers used a more compact signing space for referents that are more accessible, aligning with the trend observed in the descriptive analysis.

The fixed effect of RE [FARE in the signing space use model was significant ($\beta = -0.02$, CI: -0.05 – 0.01, $p = .277$). Although [FARE] is signed with longer duration, it is signed in a narrower signing space compared to KEDI. As mentioned before, there are different types of movement patterns associated with each expression. This could have resulted in a kinematic difference between each referent type.

In addition to linear polynomials, higher-order polynomial terms were introduced into the model to capture potential non-linear relationships between accessibility scores and the dependent variables. For the duration of REs, the cubic polynomial term for accessibility was significant ($\beta = 0.21$, CI: 0.06 – 0.36, $p = .005$). This indicates that the relationship between accessibility and duration exhibits two turning points. In other words, the model predicted that duration is sensitive to three levels of discourse with two turning points (one for maintenance and another for re-introduction).

In contrast, for the duration-normalized accumulated hand distance, the quadratic term was significant ($\beta = 0.17$, CI: 0.01 – 0.32, $p = .037$), suggesting a single turning point in the relationship between accessibility and hand movement.

The cubic term, however, was not significant ($\beta = 0.07$, CI: -0.09 – 0.22, $p = .398$), indicating that a simpler non-linear model is sufficient to describe the variation in hand distance. This finding implies that the kinematic measure hand distance is sensitive to two levels of discourse (first versus subsequent mentions). Furthermore, a significant 7th-order polynomial term ($\beta = 0.15$, CI: 0.06 – 0.24, $p = .001$) was observed in duration, which suggests even greater complexity in the relationship.

For the duration-normalized sign space use, the quadratic term was again significant ($\beta = 0.17$, CI: 0.03 – 0.30, $p = .014$), suggesting a non-linear relationship with one turning point. Neither the cubic ($\beta = 0.06$, CI: -0.08 – 0.19, $p = .403$) nor higher-order terms contributed significantly to this model, indicating that the reduction in signing space as accessibility increases can largely be explained by a simpler non-linear trend. Similar to hand distance, this results could mean that the kinematic measure sign space use mostly distinguishes between first and subsequent mentions.

5.3 Event analysis

5.3.1 Statistical data analysis

Subsequent analyses of the log-normality checks indicated that the three phonetic measures—duration, accumulated hand distance, and sign space use—deviated from normal distribution (duration: $W = .83, p < .001, D = .79, p < .001$; accumulated hand distance: $W = .97, p < .001, D = .66, p < .001$; sign space use: $W = .97$, $p < .001, D = .52, p < .001$). As in the previous analysis, all phonetic measures were log-transformed following the guidelines of Feng et al. (2014). Additionally, data points that exceeded ± 3 times the interquartile range were considered outliers and removed, accounting for 3% of the total data. After these adjustments, the data more closely followed a normal distribution (duration: $W = .99, p = .27, D = .49$, $p < .001$; accumulated hand distance: $W = .99, p = .12, D = .25, p < .001$; sign space use: $W = .97, p < .001, D = .75, p < .001$).

Transformed phonetic measures were conducted using mixed-effects models implemented with the `lmer` function from the `lme4` package (Bates et al., 2024).

5.3.2 Results

Using MediaPipe's pose landmark estimation, I analyzed 42724 frames across nine narrative events. Narrative 10, Event 5 represented the largest share, accounting for 20.6% (8811 frames), followed by Narrative 3, Event 4 with 16.3% (6962 frames). Narrative 2, Event 6 contributed 10.3% (4403 frames), and Narrative 3, Event 2 comprised 11.0% (4705 frames). Other notable proportions include Narrative 9, Event 4 (9.22%, 3940 frames), Narrative 6, Event 8 (11.6%, 4977 frames), Narrative 2, Event 5 (7.69%, 3286 frames), Narrative 6, Event 4 (7.86%, 3358 frames), and Narrative 2, Event 7 (5.34%, 2282 frames).

The normalized phonetic values were categorized by participant, event, and stimuli, yielding 233 unique data points for statistical analysis (Narrative 10, Event 5: 28, 12%; Narrative 2, Event 5: 27, 11.6%; Narrative 6, Event 8: 27, 11.6%; Narrative 2, Event 6: 26, 11.2%; Narrative 3, Event 4: 26, 11.2%; Narrative 9, Event 4: 25, 10.7%; Narrative 6, Event 4: 25, 10.7%; Narrative 2, Event 7: 25, 10.7%; Narrative 3, Event 2: 24, 10.3%). The phonetic measures (hand distance and sign space use) were log-transformed and averaged across both the left and right sides for subsequent analysis.

Figure 18 presents the results for RE duration (given in seconds) by Group.

It can be seen that native signers have a shorter average duration ($M = 3.21$ seconds, $SE = .223$) than late signers ($M = 4.12$ seconds, $SE = .269$).

Figure 19 illustrates the average duration-normalized hand distance by group.

Native signers exhibited only slightly higher mean normalized amplitude ($M = 1.08$, $SE = 0.0315$) compared to the late signing group ($M = 1.02$, $SE = .0341$).

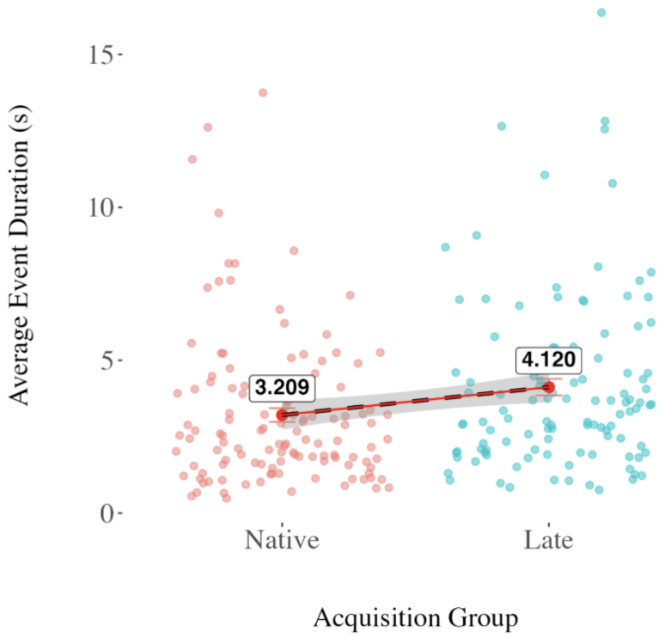


Figure 18. Average Event Duration (s) by Group. In the plots, each dot is a single observation (RE). The red line and points represent the mean observed value for each bin. The dashed black line is a regression line fitted across the combined data from all accessibility bins

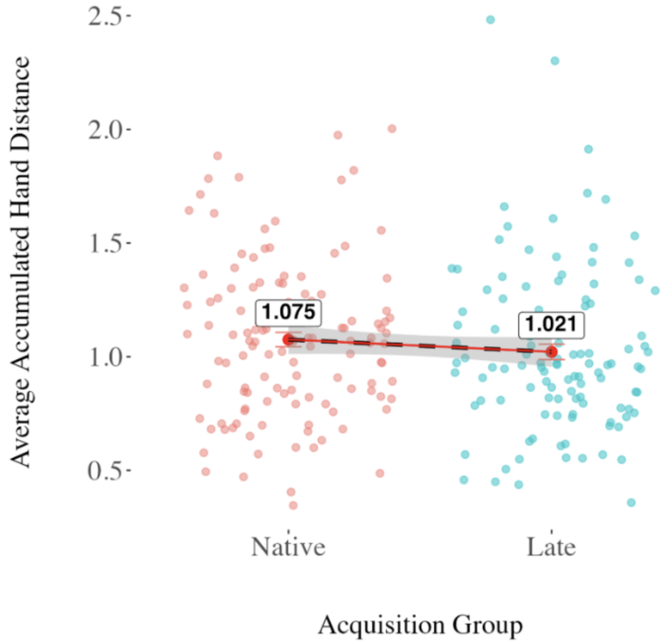


Figure 19. Average Duration-Normalized Hand Distance by Group. In the plots, each dot is a single observation (RE). The red line and points represent the mean observed value for each bin. The dashed black line is a regression line fitted across the combined data from all accessibility bins

Lastly, Figure 20 presents the mean sign space use, which measures the mean Euclidean distance between the wrists and middle shoulder point, normalized by duration.

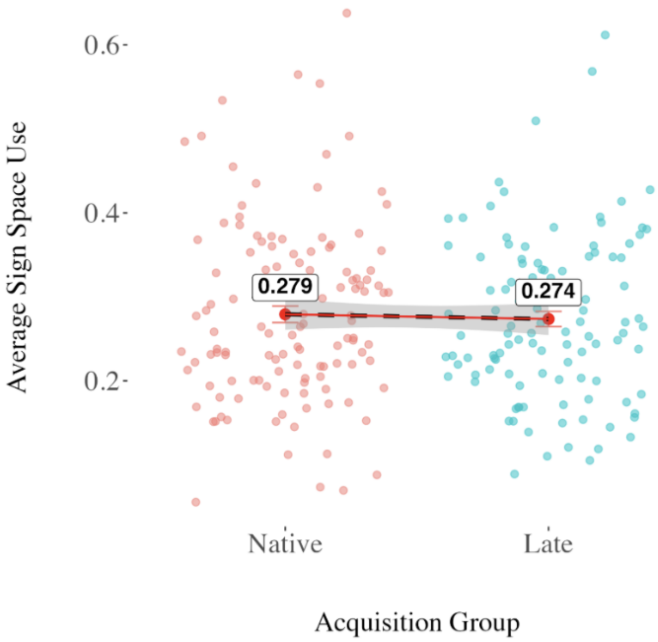


Figure 20. Average Duration-Normalized Sign Space Use by Group. In the plots, each dot is a single observation (RE). The red line and points represent the mean observed value for each bin. The dashed black line is a regression line fitted across the combined data from all accessibility bins

Native signers again demonstrated slightly higher mean normalized sign space use ($M = .279$, $SE = .00981$) than late signers ($M = .274$, $SE = .00908$).

I fit three mixed effects models to analyze the data. For the utterance-level analysis, the models incorporated Group (Native, Late) as a fixed factor, with the phonetic measures (Duration, Duration-Normalized Accumulated Hand Distance, and Duration-Normalized Sign Space Use) as the dependent variables.

(15) $\text{Model}_{\text{Duration}}$ = lmer(log(Duration) Group + (1 | Participant) + (1|Event) + (1 | Stimuli))

(16) $\text{Model}_{\text{Hand.Distance}}$ = lmer(log(Distance) Group + (1 | Participant) + (1|Event) + (1 | Stimuli))

(17) ($\text{Model}_{\text{Sign.Space.Use}}$ = lmer(log(Space) Group + (1 | Participant) + (1|Event) + (1 | Stimuli))

Random effects were specified for Participant, Stimuli Number, and Event Number. The model formulas are given in (46-48).

The mixed-effects regression models of the utterance-level analysis are presented in Table 8.

Table 8. Model results for utterance-level analysis by Group

	Duration (log)			Duration-Normalized Accumulated Hand Distance (log)			Duration-Normalized Sign Space Use (log)		
Predictors	Estimate	CI	p	Estimate	CI	p	Estimate	CI	p
Group [Native]	-0.14	-0.26 – -0.02	0.020	0.02	-0.07 – 0.11	0.723	-0.01	-0.12 – 0.10	0.885
Random Effects									
σ^2	0.20			0.06			0.07		
τ_{00} Participant	0.07			0.05			0.08		
τ_{00} Event	0.08			0.00			0.02		
τ_{00} Narrative	0.19			0.01			0.01		
ICC	0.63			0.49			0.64		
N Participant	28			28			28		
N Event	6			6			6		
N Narrative	5			5			5		
Observations	225			225			225		
Marginal R^2 / Conditional R^2	0.035 / 0.642			0.002 / 0.492			0.000 / 0.636		

Accordingly, duration was the only significant measure influenced by the acquisition group. Specifically, native signers produced utterances with shorter durations compared to late signers ($\beta = -0.14$, CI: -0.26 – -0.02, $p = 0.020$), indicating more temporal efficiency among natives. More specifically, this finding aligns with Braem (1999), suggesting that native signers may use duration as an economical linguistic strategy during event retellings, possibly avoiding redundancy.

In contrast, the models for duration-normalized accumulated hand distance and sign space use did not show significant differences between groups ($p > 0.7$). Native and late signers exhibited similar average hand movement and signing space use in their signing, suggesting that the acquisition group does not influence these specific phonetic or kinematic measures at the utterance level. These results imply that while temporal efficiency (duration) varies by acquisition group, hand movement patterns, and spatial signing behaviors are not significantly impacted by age of language exposure.

As to why this occurs, it is possible to suggest language conventionalization. In established sign languages like TİD, conventionalized norms govern phonological parameters like the amount of movement and signing space size. These norms are typically acquired by all signers, regardless of age at acquisition, leading to uniformity in these phonetic features. This contrasts with emerging sign languages, such as Nicaraguan Sign Language, where younger signers have been observed to use more restricted signing spaces, reflecting ongoing language evolution (Flaherty et al., 2023).

As an interim summary, the results of the phonetic analyses revealed patterns across multiple levels, including discourse contexts, referent accessibility, and acquisition groups (native vs. late signers).

For discourse contexts, the analyses indicated that introducing new referents required significantly greater phonetic effort compared to maintaining or re-introducing them. Specifically, referent first mentions exhibited the longest durations, largest accumulated hand distances, and widest sign space use. Conversely, referent maintenance resulted in shorter durations and narrower sign space use, though its effects on hand movement amplitude were less pronounced.

When analyzing phonetic measures in relation to referent accessibility, linear analysis showed that higher accessibility scores were associated with reduced phonetic effort. More accessible referents were articulated with shorter durations, smaller hand movements, and narrower sign space, indicating reduced physical and cognitive effort during production. In addition, higher-order polynomial analysis demonstrated that duration as a phonetic measure seems to be sensitive to three levels of discourse as the cubic polynomial term was significant in the model. In contrast, the kinematic measures, namely hand distance and signing space use, displayed binary sensitivity to discourse since the quadratic term was significant in these models and the cubic was not.

Finally, for the utterance-level analysis, the results revealed that native signers demonstrated shorter utterance durations compared to late signers. However, no notable differences were found between native and late signers in accumulated hand distance or sign space use. These results imply that native signers' earlier exposure to sign language may lead to greater temporal efficiency or economicity but do not significantly affect spatial or kinematic measures.

Overall, these findings suggest that cognitive, linguistic, and motor processes work in tandem to optimize efficiency while accommodating communicative demands. These results will be explored and elaborated on further in Chapter 6, where their implications for referential accessibility, linguistic efficiency, and acquisition will be discussed in detail.

CHAPTER 6

DISCUSSION

This chapter presents a general discussion on the findings related to discourse cohesion and phonetic measures in TİD narratives. It also examines how these results contribute to our understanding of language economy, accessibility, and pragmatic and phonetic strategies employed by native and late signers.

6.1 Discussion of discourse cohesion and referential accessibility

The questions that the present study examined with regard to discourse cohesion and referential accessibility in TİD narrative production were as follows:

(1) How do discourse context and type of referring expression relate to the accessibility scores of referents?

(2) Do native signers track accessible referents more than late signers, which is realized as minimizing the distance between current and previous mentions, tracking more topically salient referents, and engaging in less referential competition?

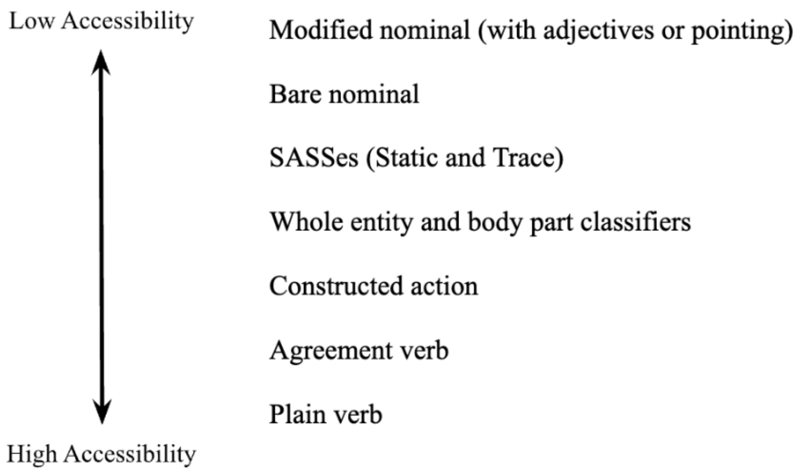
The prediction was that both discourse context and RE type could be strongly correlated with accessibility, given the theories of linguistic efficiency and referential accessibility. Zipf's (1949) Principle of Least Effort, Grice's (1975) Maxim of Quantity, and Sperber and Wilson's (1986) Relevance Theory all predict that we use phonetically minimal or reduced forms when possible. Ariel's (1990) accessibility hierarchy also predicts low accessibility markers, such as full noun phrases, to be used for less accessible referents, while high accessibility markers, such as pronouns, to be employed for highly accessible or previously mentioned referents. Given the phonetic forms associated with these nominals, Arial's hierarchy is closely in line with the aforementioned linguistic efficiency and cooperation principles. Note that all of these proposals were made with spoken language in mind, but their predictions also extend to signed discourse. Nominals were generally used to introduce less accessible referents, while more reduced forms like constructed action (CA) and verbs tracked highly accessible referents in ASL and BSL, among others (Frederiksen & Mayberry, 2016; Morgan, 2000; Pizzuto, Rossini, Sallandre, & Wilkinson, 2006; Swabey, 2002). The results of the present study on TİD data presented are, overall, in line with previous research. I present a detailed discussion of the findings below.

The results showed that both discourse context and the RE types (NOM, CL, CA, VERB) successfully predicted the accessibility scores of referents in TİD narratives. The model that included Discourse as a predictor along with Group had an R2 value of 0.877, which means that approximately 88% of the variance in accessibility scores was explained by Discourse. To compare, the model with RE Type and Group as predictors had an R2 value of 0.521, explaining about 52% of the variance in accessibility scores. This means that although both factors significantly predicted the outcome variable, Discourse explained the distribution of the accessibility scores better than the RE Type.

As to the observed relationship between accessibility and RE type, it is important to reiterate that this thesis is the first to adopt a graded accessibility scale to TİD. The mixed effects models demonstrated that nominals are predominantly used to introduce referents with low accessibility, while classifiers and constructed actions are utilized for referents with moderate accessibility. Verbal constructions, such as agreement and plain verbs, are primarily employed to maintain highly accessible referents. Furthermore, post-hoc pairwise comparisons revealed that each RE level significantly differed from the others in terms of referential accessibility. Although these findings are predictable given the literature on linguistic efficiency and discourse cohesion in both spoken and sign languages, one important methodological contribution of this study is the adaptation and validation of Toole's (1996) graded accessibility scale in the context of TİD. This quantitative approach, therefore, enabled a finer-grained understanding of the relationship between referential forms and their discourse functions, which was previously unexplored for TİD. As a result, it became possible to propose a revised hierarchy of accessibility for TİD.

We had proposed a hierarchy of referential accessibility for TİD narratives in Keleş et al. (2023) as in (3). However, the original hierarchy for TİD did not quantize accessibility scores, i.e., they were based on the frequency of occurrence of RE forms. Given the distribution of these RE types and the mean accessibility score of the referent that they track (i.e., the degree of accessibility a referent has in the mind of the addressee given certain pragmatic factors like distance, unity, salience, and competition) in the present study, the accessibility hierarchy for TİD can be revised as in (4)[12].

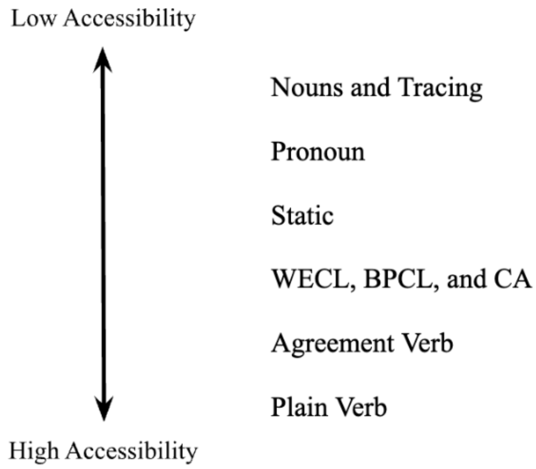
(3) Keleş et al.'s proposed accessibility hierarchy in TİD.



(Keleş et al., 2023, p. 27).

[12]In Keleş et al., we did not analyze static and tracing tools separately but we had grouped them under one category and argued that they were higher than nominals but lower than classifiers on the accessibility hierarchy. This thesis analyzed the two constructions separately, and the findings suggest that static and tracing tools may occupy distinct positions on the hierarchy if we consider the mean accessibility scores of their tracked referents. Specifically, static tools were associated with referents approximately twice as accessible as those tracked by tracing tools. Another notable difference between the two proposed hierarchies pertains to the position of CA. In the former hierarchy, CA, which was grouped with verbal constructions under Zero Anaphora, was assumed to rank higher than classifiers. However, I have shown in the present study that CA and classifiers (WECL and BPCL) function quite similarly with respect to the mean accessibility scores. The production experiment also demonstrated that verbs were mostly utilized for maintenance and the primary re-introduction forms were nominals, classifiers, and CA.

(4) Present study's revised accessibility hierarchy for TİD.



Taken together, this shows that referential accessibility alone cannot justify the choice of CA over a classifier in TİD, and other pragmatic factors might be at work. I suspect that signers might prefer CA over CL, regardless of the fact that they track similarly accessible referents, because of the following reason: it is possible to use the spatial affordances of the signing space and engage in concurrent forms, which refer to the simultaneous signing of multiple and distinct REs (Cormier, Smith, & Zwets, 2013; Czubek, 2017; Dudis, 2004; Janzen, 2004; Liddell, 2003; Smith & Cormier, 2014). Thanks to the affordance of simultaneous expression, signers can track one referent from the character perspective (i.e., begins enacting the protagonist) through CA, and with the hands, they can track one or more other distinct referents from the observer perspective by employing classifiers. For this reason, when the context necessitates tracking multiple referents at the same time, employing CA together with CL can be more economical.

Another important observation was that TİD signers in this study used pronouns very sporadically (40 in total), and the observed occurrences were mostly used to re-introduce referents. This finding parallels that of other studies on ASL (Czubek, 2017; Frederiksen & Mayberry, 2016; Swabey, 2002) and Auslan (Hodge, Ferrara, & Anible, 2019), which also suggested limited use of pronominal references in both simple and more complex narratives. In contrast, some other research indicated that the signers of LSC (Bel, Ortells, & Morgan, 2015), BSL (Morgan, 2005), and DGS (Perniss & Özyürek, 2015) frequently employ overt pronouns (up to %30), especially to maintain a previous referent in the narrative discourse. To explain the low occurrence of pronouns in their ASL data, Czubek also proposed that genre can affect the frequency of such referents. This is because Czubek found that personal narratives use more explicit pronouns than picture or video retellings, likely due to the need for clarity in personal storytelling. Additionally, Frederiksen and Mayberry reported a preference for pronominals when referring to objects rather than subjects in ASL. If this is the case, factors such as genre dependency and object preference may explain the low distribution of pronouns in our data since we included only one genre (narrative retelling) and considered only subject referents.

I had also predicted that delayed or late exposure to the first language might affect discourse cohesion and accessibility, considering our earlier reference tracking work (Keleş, Atmaca, & Gökgöz, 2023) and the literature on the age of acquisition (Boudreault & Mayberry, 2006; Cheng & Mayberry, 2021; Kayabaşı & Gökgöz, 2023; Sevgi & Gökgöz, 2023). In our previous study, we had found that nativeness significantly affected the production of zero or null forms, including CA, PLAIN, and AGR (Keleş et al., 2023). In other words, late signers were more overexplicit in their reference tracking, and they produced fewer implicit forms than natives, especially for reference maintenance and re-introduction. Considering this result, I predicted that native signers would refer to more accessible referents overall in their narratives than late signers. In this sense, a group's having a higher mean of accessibility implies that the signers in that group are continuously tracking the same entity without frequent shifts to other competing referents. This implication follows from how the accessibility score measured in this thesis is defined.

The results of the production experiment and accessibility scores supported this prediction. Acquisition had a significant effect on accessibility scores. The referents tracked by native signers had more mean accessibility for most RE types than the late signing group. There were both similarities and differences in the referential strategies for RE usage by either group. Both groups used nouns as a primary strategy to introduce and re-introduce referents. On similar grounds, both natives and late signers used CA and verbal constructions for reference maintenance. These strategies can be explained by the principles of abovementioned linguistic efficiency and referential accessibility, which is not surprising.

The two groups also diverged in the use of these REs. For introduction, late signers favored TRACE more than natives. For maintenance, native signers produced more CA, PLAIN, and AGR than late signers, who, in turn, used twice as many NOMs to maintain the previous referent. The statistical model of Group and RE Type in the present study also revealed an interaction effect between being a native signer and using a nominal. In that case, NOM reduced the accessibility score more when native signers used it. This suggests that late signers must be overusing nominals for highly accessible referents. With the graded accessibility analysis, it is also possible to see that late signers diverged by using tools like BPCL, CA, and PLAIN to introduce more accessible referents, while native signers either avoided these tools or used them for less accessible ones in the context of introduction. Despite this, both groups showed similar patterns in the maintenance and re-introduction contexts, tracking highly accessible referents, though native signers sometimes tracked slightly more accessible ones. Statistical analyses also confirmed that native signers had significantly higher accessibility scores overall.

The increased accessibility scores for native signers can be understood considering the syntax-discourse interface (Sorace, 2011; Sorace & Filiaci, 2006). Since late signers' tracking referents with lower mean accessibility indirectly refers to more topic shifts and less reliable use of zero forms (such as CA and verbs), it is possible to suggest that late signers do not behave economically in terms of discourse cohesion since they may struggle with this interface. The syntax-discourse interface, which involves the integration of syntactic structures with discourse-level considerations, is a known area of vulnerability for non-native speakers (Sorace, 2011). This vulnerability has been consistently observed among second language learners in both spoken (Gullberg, 2006; Tomlin, 1987; Williams, 1988) and sign language acquisition (Bel et al., 2015; Frederiksen & Mayberry, 2019). Although only a few studies investigated and reported over-explicitness by signers with first language delay (Becker, 2009; Swabey, 2002), we could extend these findings by suggesting that late signers face particular challenges in managing this interface effectively. According to Sorace's Interface Hypothesis 2011, interfaces between different cognitive domains, particularly between syntax and discourse pragmatics, are especially vulnerable in late language acquisition. This vulnerability stems from the increased cognitive load required to simultaneously coordinate syntactic knowledge with discourse-pragmatic constraints in real-time language processing. For late signers, this means they must consciously monitor and integrate both syntactic rules for referring expressions and discourse-level information about referent accessibility, a process that remains cognitively demanding even at advanced levels of proficiency. This, in turn, can impact their ability to select appropriate referring expressions and maintain discourse cohesion in a native-like manner. Developing robust automaticity at the syntax-discourse interface might allow native signers to use more reduced forms, such as pronouns or zero anaphora, and consequently track referents with higher accessibility compared to late signers.

In this respect, the present study aligns well with previous observations (Becker, 2009, for DGS signing children; Cormier et al., 2013, for BSL; Gür & Sümer, 2022; Gür, 2024; and Keleş et al., 2023 for TİD). Becker (2009), for instance, found that late-signing children for DGS lagged behind their native peers in terms of their narrative abilities. In Becker's study, late-signing children struggled significantly with maintaining and shifting references consistently, suggesting that they had trouble keeping track of and coherently managing the narrative flow compared to their native-signing peers. The authors concluded that late-signing children were less proficient in terms of using the spatial affordances of the signing space and maintaining reference than natives due to the lack of frequent exposure to sign language. Similarly, Cormier et al. (2013) reported similar findings for BSL. While native signers mostly used CA without a nominal (e.g., explicit) reference to maintain a referent, early and late signers used zero anaphora to a lesser extent, similar to our results. In the present study, it was also found that native signers tracked more accessible referents and used more implicit forms compared to late signers. This indicated lower occurrence of topic shifts and less referential competition by native signers than late signers.

It should be noted, however, that the results are not fully in line with previous research. Gür and Sümer (2022), for instance, found that delayed first language acquisition did not impede late signers' referent introductions in TİD, unlike the results of this thesis. Gür and Sümer, however, examined the age of the participants (i.e., children versus adults) and not the age of acquisition (i.e., native versus late). Gür's follow-up study (2024) examined late and native signers' introduction of inanimate referents. It also concluded that late acquisition does not hinder referent introductions. One reason that those studies did not find an age of acquisition effect might be because they only examined referent introductions and excluded other discourse contexts. The present study examined all three discourse contexts (introduced, maintained, and re-introduced) and found that late signers diverged from native signers in their referent maintenance and re-introduction strategies, which affected the overall accessibility of referents.

To my knowledge, no previous research on sign language has quantified referential accessibility using a tool such as the one by Toole (1996) and explored how the age of acquisition impacts variations in the accessibility levels of referents across different contexts. This is important because quantifying referential accessibility allowed for a more nuanced understanding of how different factors, such as the age of acquisition, affect the efficiency of discourse cohesion in sign languages and, more specifically, in TİD. Using the graded accessibility, it was possible to see how being a native and late signer interacted with the use of RE forms. It was observed that native signers used nominals more frequently to decrease accessibility compared to late signers. This indicates that native signers are better able to modulate the use of explicit referring expressions in accordance with accessibility scores.

Furthermore, CLs were found to be less effective as accessibility-decreasing tools among native signers compared to late signers. This suggests that native signers have a higher threshold for employing CLs, meaning that they require referents to be more accessible before using this type of referring expression. Both observations provide further support that first-language delay affects discourse cohesion, possibly due to sensitivity to the syntax-discourse interface (Sorace & Filiaci, 2006). Such insights can then inform educational practices, aiding in the development of targeted interventions to support communicative competence in late signers.

6.2 Discussion of phonetic reduction and referential accessibility

The research questions regarding phonetics and discourse were as follows:

(5) Do phonetic measures like sign duration, hand distance, and sign space use correlate with referential accessibility and discourse context in TİD narratives?

(6) Are such phonetic measures in the re-telling of events in signed narratives age-sensitive? In other words, does delayed or late exposure to first language affect sign duration, hand distance, and sign space use during the narration of an event in a story among deaf TİD signers?

For the first question, it was predicted that the articulatory phonetics of nominal REs in TİD (sign duration, hand distance, and sign space use) would be in line with their referential accessibility and discursive context. That is, a nominal RE tracking a highly accessible referent was predicted to be signed faster, include less hand movement, and uttered in a narrower sign space compared to an RE tracking a less accessible referent. This prediction followed from theories of efficiency (Grice, 1975; Sperber & Wilson, 1986; Zipf, 1949), referential accessibility (Ariel, 1990; Givon, 1983; Gundel, Hedberg, & Zacharski, 1993), and studies of phonetic reduction in sign languages (Flaherty, Sato, & Kirby, 2023; Hoetjes, Krahmer, & Swerts, 2014; Mauk & Tyrone, 2012; Stamp, Dachkovsky, Hel-Or, Cohe, & Sandler, 2024; Tyrone & Mauk, 2010, 2012). Overall, the results of the phonetic analyses on the nominal REs (FARE 'mouse' and KEDI 'cat') were consistent with the predictions. The results showed that introduced REs had the longest duration, followed respectively by re-introduced and maintained ones. Predictable contexts (i.e., maintenance and re-introduction) that included already old or already established referents mean durations similar to each other.

The kinematic analysis of the narrative data furthermore demonstrated that the remaining phonetic measures, namely accumulated hand distance and signing space use, changed significantly as a function of both discourse context and accessibility. In other words, as a referent's accessibility increased (i.e. when the context rendered the referent more predictable), the amount of hand movement decreased, and the nominal RE was signed in a narrower sign space. Overall, this finding adds empirical evidence to the observation that accessibility is a significant predictor of phonetic reduction in TİD. In terms of the discourse contexts, introduction exhibited a significant reduction compared to the grand mean across all phonetic measures. In contrast, maintenance achieved statistical significance only for duration reduction. The significant effect of phonetic reduction in this discourse context was not observed for other kinematic measures, such as hand distance and sign space use. Since maintenance reached statistical significance only for the duration and was not significant in other measures, we can suggest that the level of phonetic reduction for maintenance and re-introduction was relatively similar. The ordinal accessibility analysis also revealed a non-linear relationship between accessibility and phonetic measures, with duration showing sensitivity to three discourse levels and two turning points, while hand distance and sign space use primarily distinguished between first and subsequent mentions with simpler non-linear trends.

Several conclusions can be drawn from this result. First, the degree of phonetic reduction in TİD, a conventionalized sign language, increases as the referent's discourse status becomes predictable or its accessibility increases, indicating that signers economize their articulatory effort in an appropriate pragmatic context. This is not surprising given how TİD signers adhere to the principles of linguistic efficiency (Zipf, 1949) and cohesion (Ariel, 1990; Givon, 1983; Gundel et al., 1993).

Secondly, as revealed by the graded accessibility analysis, that kinematic measures were sensitive to a binary distinction between new versus old/predictable information. The duration, on the other hand, differed for introduction, maintenance, and re-introduction. Here, we can infer that TİD signers are particularly attuned to the different discourse contexts when modulating duration, suggesting that the temporal aspect of sign production serves as a more finely-tuned marker of discourse cohesion than kinematic aspects, which only marks new versus predictable in TİD. This might be because tuning duration might require less physical effort than the kinematic parameters for each discourse level.[13]

To my knowledge, this analysis is the first computer vision work employed on TİD narratives and it is the first computational analysis that examined the relationship between graded referential accessibility and kinematic measures in sign language. Given the results, it is clear that this methodology is particularly valuable and appropriate for investigating the interplay between physical articulation and discourse-level accessibility in the visual-spatial modality. Comparing the results to previous work, it can be seen that these findings mostly supported Stamp, Dachkovsky, et al.'s (2024) findings on Israeli Sign Language (ISL) narratives, which also showed that nominal REs for introduction had the longest duration, followed by re-introduction and maintenance. They also reported similar findings in other phonetic measures like distance, speed, variance, volume, and a measure of the sign space for three different articulators (hand, torso, and head). Their conclusion was that predictable referents (i.e., those with high accessibility) underwent phonetic reduction.

However, the results of the remaining phonetic measures reported in this thesis (hand distance and sign space use) are not comparable to other research since I did not use a motion tracker camera. Instead, MediaPipe, an open-source markerless

[13] As a result, we only measured hands in this analysis, and other body parts (e.g., head and torso) might have a different pragmatic function, which requires future investigation.

pose estimation library, was used to obtain these values, and the output does not correlate to centimeters or meters (as reported in Stamp, Dachkovsky, et al., 2024) because the library provides coordinates that are normalized between 0 and 1, which represent the relative position within the frame. Therefore, the measurements described in Chapter 5 are dependent on the resolution and scaling of the video rather than absolute physical measurements.

As to the observation that kinematic parameters were sensitive to new versus predictable distinction, our results were mostly consistent with Stamp, Dachkovsky, et al. (2024), who reported the same sensitivity for all their measures. I had previously found significant differences among all levels of discourse contexts for referential accessibility in Chapter 5, where the discourse cohesion results were reported. In other words, the results of both this thesis and our earlier work (Keleş et al., 2023) significantly alluded to the following hierarchy in terms of overt marking and accessibility: introduction > re-introduction > maintenance.

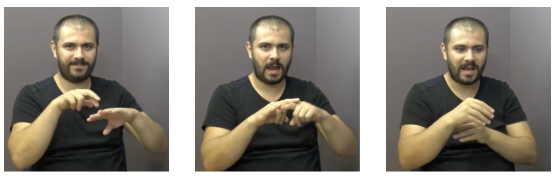
Perniss et al. (2015), for instance, compared spoken German, co-speech gesture, and DGS and reported a tripartite difference in overt marking, similar to the one described above. However, other research, like Frederiksen and Mayberry (2016), did not find a statistical difference between maintenance and re-introduction. Similarly, our reduction analysis also showed that TİD signers are mainly sensitive to new versus predictable (maintenance and re-introduction) information in terms of phonetic marking, with the exception of duration.

When we examine accessibility scores discretely, though, it is possible to make some interesting observations. For the duration results, three levels of accessibility are worth mentioning, as can be seen in Figure 21, which consists of the reproduced plots reported in Chapter 5. While the other levels captured the linear trend, the durations of REs with scores at the extreme ends (-2 and 5) slightly diverged from my expectations. This discrepancy, however, did not spread to other kinematic measures (hand movement and sign space use). In these measures, the observed linear trend was more preserved.

I had the general prediction that the least accessible REs (with a score of -2) would have the longest duration and the most accessible REs (with a score of 5) would have the shortest. One possible explanation for this could be attributed to the lack of sufficient data for these levels, with 23 observations for level -2 and 25 for level 5, which might have led to this variation. Moreover, for a referent to be assigned a score of -2, there needs to be at least two other competing referents preceding it. As a result, by the definition of the accessibility scale that I used in this thesis, a referent obtaining a score of -2 implies that it can be the third introduction in a sequence, following two other introductions. An example of this is given in (7). In this example, all of these nominals constitute the first three REs signed subsequently in this signer's 6th narrative. This means that this signer had produced 5 other independent narratives prior to this, all of which took place in the canon of Tom and Jerry.

The third referent, MOUSE, received a distance and unity score of 0 because it was the first mention of Jerry in this narrative, and also, it was not mentioned in the immediately previous narrative either (in this case, it would have gotten a score of 1). Additionally, we can see that its competition score is -2 since two other introductions precede it. Put differently, the signer had already started setting the scene before mentioning Jerry by introducing other referents.

(7)



STAIRS CARPET MOUSE

Discourse Context:	Introduction	Introduction	Introduction
Distance and Unity Score:	0	0	0
Topicality Score:	0	0	0
Competition Score:	0	-1	-2
Total Accessibility Score:	0	-1	-2

I suspect that the duration of such referents might be reduced compared to other introductions that received a 0 or 1 due to contextual familiarity. That is, by the time a third referent is introduced, the signer may assume that the addressee is now more

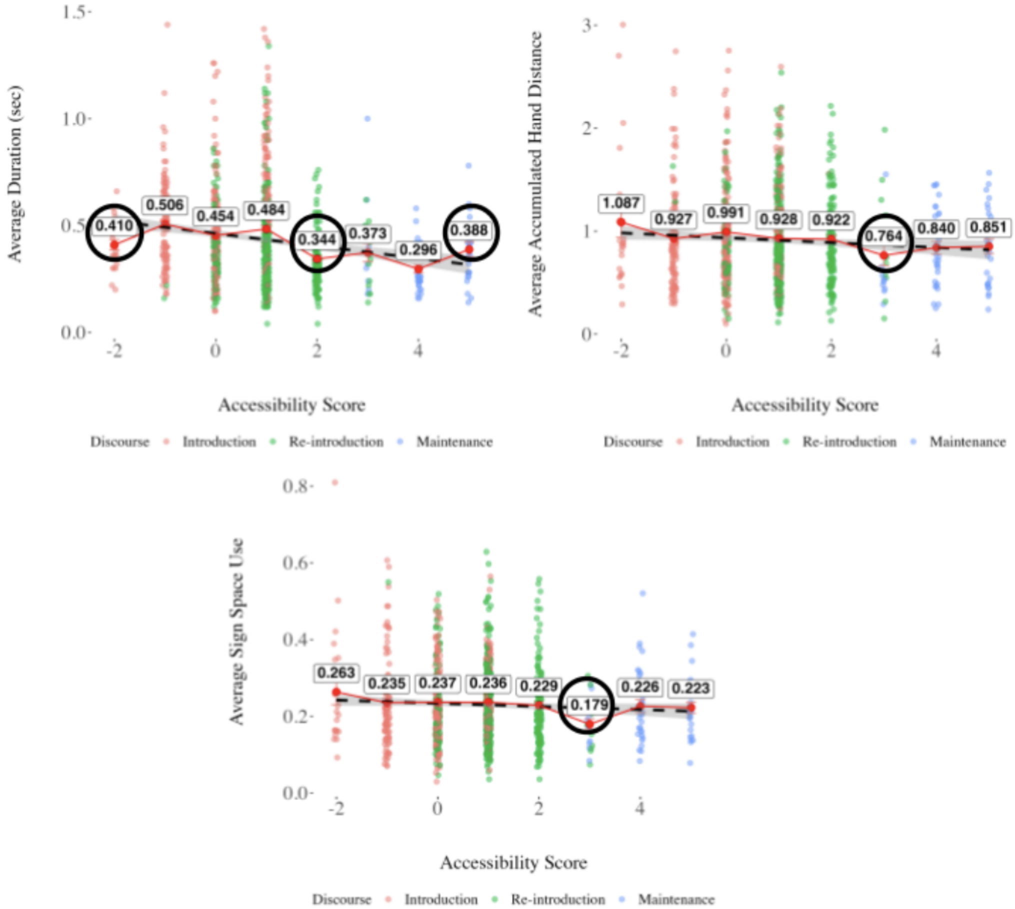


Figure 21. Average RE Duration, Hand Distance, and Sign Space Use by Accessibility Score. Each dot is a single observation (RE). The red line and points represent the mean observed value for each bin. The colors represent the discourse contexts. The dashed black line is a regression line, fitted across the combined data from all accessibility bins. The black circles have been added to emphasize certain levels of accessibility

familiar with the structure and flow of the narrative, and thus, the need to emphasize the new referent might decrease. As a result, the duration of the REs with an accessibility score of -2 could be shorter than other types of introductions. However, if this is the case, it is important to note that this type of reduction was only observed in duration and not in any of the other phonetic measures. Also, this observation could be explained with the idea of rhythmic list prosody, proposed and detailed by Erickson (1992) and Selting (2007) for spoken languages, who noted that prosodic features play a significant role in how speakers organize and present list structures. Selting (2007) posited that speakers use specific pitch movements, such as rising intonation on non-final items and falling intonation on the final item, to signal continuation and completion. Since referents with -2 are the last element of a list, item prosody might explain why the duration is lessened for such referents. This pattern of reduced duration might also be linked to findings by Brozdowski and Emmorey (2023), who noted a greater degree of rhythmic list prosody for pseudosigns by deaf signers where temporal regularity aided in organizing sequences, even in the absence of semantic meaning.

As to why there is an unexpected increase from an accessibility score of 4 to 5 (Figure 21), let us again revisit the definitions of the adopted scale from Toole (1996). For a referent to obtain a score 5, it must be at least the third continuation of the same referent in a row, resulting in zero competition and high topicality. As illustrated in Chapter 4, the most accessible referents (i.e., those who scored an accessibility of 5) were most often tracked with CA (M = 42%), followed by VERB (M = 32%), CL (M = 19%), and then NOM (M = 7%). The use of a nominal to track a highly accessible referent is not a common observation cross-linguistically, and this disobeys the aforementioned principles of linguistic efficiency (Grice, 1975; Sperber & Wilson, 1986; Zipf, 1949) and referential accessibility (Ariel, 1990; Givon, 1983; Gundel et al., 1993). Upon another inspection, it was found that, among native signers, only 4% of referents with a score of 5 belonged to nominals. This percentage increased to 12% for late signers. As a result, only 11 of the nominal data points reported for 5 scorers pertain to native signers. Considering the over-use of nominals by late signers, I suspect that acquisitional effects might have contributed to the observed variance here.

The final point that I would like to make concerns the sharp decrease (approximately 30%) in duration from an accessibility score of 1 to 2, followed by another, albeit less striking decrease (20%), from 3 to 4 (Figure 21). This is interesting because an accessibility score of 2 also marks the change from introduction to re-introduction, as can be seen by the colored points in the plots. In contrast, a change from an accessibility score of 3 to 4 indicates a shift from re-introduction to maintenance. As a result, RE duration seems to be sensitive to all three levels of discourse: once a certain accessibility threshold is satisfied, the descriptive results suggest that duration is reduced accordingly.

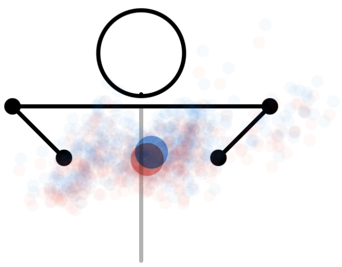
A similar drop can be observed around levels 2 to 3 for the other phonetic measures, with a 17% decrease for hand distance and a 22% decrease for sign space use (Figure 21), which might give insight into why the main effect of maintenance does not reach statistical significance in the first model. Unlike RE duration, significant reductions in these phonetic measures only become apparent at an accessibility score of 3, beyond which hand distance and sign space use stabilize. Since there is only one noticeable decline observed for accessibility scores (from 2 to 3) as opposed to duration that included two important decreases (one from 1 to 2, another from 3 to 4), it could be claimed that any reduction in kinematic features is mostly based on a two-way distinction between new versus predictable referents. This empirical observation is parallel to those reported in Stamp, Dachkovsky, et al.(2024) for ISL. However, it seems that the duration parameter can be best accounted for with a tripartite classification of discourse.

As a summary, the results illustrated that phonetic measures like the duration of a sign, as well as certain kinematic measures like hand distance and sign space use, undergo reduction as a function of referential accessibility in TİD narratives. This finding mainly supports previous observations of reduction both in spoken and sign languages (Hoetjes et al., 2014; Tily & Piantadosi, 2009), all of which suggest that the quantity of form depends on three main discourse statuses: introduction, maintenance, and re-introduction. However, I found that the reduction of some phonetic measures (hand distance and sign space use) is primarily sensitive to a more binary classification of new versus predicted information. When the information status of a referent changes from new to old (i.e., already established), there appears to be a significant difference in these measures.

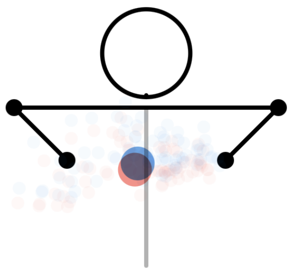
Although the findings support a strong relationship between referential accessibility and phonetic reduction, it is important to acknowledge that other phonological and prosodic factors may contribute to the observed variations in sign duration and movement. In TİD, as in other sign languages, the position of signs within sentences can significantly influence their phonetic realization. The annotation guideline used in this study, following previous work (Frederiksen & Mayberry, 2019), coded only subjects for maintenance and re-introduction contexts, which correlate with sentence-initial positions in TİD, while referents in both subject and object position were coded for introduction. As a result, introduced signs have more positional variation and this might affect the duration and kinematic parameters. Also, topic positions, which frequently align with maintenance and re-introduction functions, are often prosodically marked through specific non-manual markers like eyebrow and head movements, and are typically followed by a slight pause before the rest of the sentence (Calderone, 2020; Keleş & Gökgöz, 2022; Nespor & Sandler, 1999). Although this pause is generally less pronounced than sentence-final pauses, it could nevertheless influence the duration of target signs.

Moreover, phrase-final lengthening is a well-documented phenomenon in both spoken (Edwards, Beckman, & Fletcher, 1991) and signed languages (Wilbur, Malaia, & Shay, 2012). Studies in Quebec Sign Language and Sign Language of the Netherlands have reported strategies for phrase-final marking, including longer final holds and increased spatial displacement (Crasborn, Van der Kooij, & Ros, 2012). Additionally, kinematic studies in American Sign Language (ASL) suggest that final lengthening involves both temporal and spatial strategies, such as longer sign holds and release movements (Stewart, 2014; Tyrone & Mauk, 2010). However, its influence on RE duration and kinematics in the context of the present study may be mitigated since the REs under investigation, such as subjects and objects, precede the verb and are thus positioned outside the phrase-final environment where lengthening effects are most prominent. This suggests that target signs, whether functioning as subjects or objects, are less likely to undergo phrase-final lengthening effects.

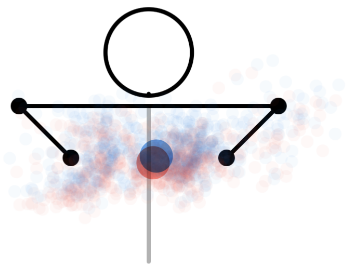
It is also important to acknowledge that the surrounding signs and the overall rhythmic structure of the sentence may still influence the prosodic realization and position of the target signs. However, to minimize the effects of the position of the surrounding signs, the boundaries of the target sign were identified following previous literature on the selected finger constraint (Sandler, 1999). The sign onset was determined by the transition of selected fingers from the previous sign to the target sign. Figure 22 illustrates the average positions of the hands in the FARE and KEDI signs (collapsed) produced by the TİD signers. While maintained nominals are mostly signed proximal to the signer, we can see more positional variation for introduced and re-introduced referents, some of which were signed distal to the signer (e.g., in the upper left side of the signer).



(a) Introduced nominals



(b) Maintained nominals



(c) Re-introduced nominals

Figure 22. Average right (red) and left (blue) hand positions for (a) introduced, (b) maintained, and (c) re-introduced nominals FARE and KEDI (collapsed). The larger points in each plot represent the grand mean for right and left hands.

Although the scope of this thesis does not include co-articulatory effects, during annotation I noticed that the some distal localizations of the target sign (e.g., to the right of the signer) illustrated in Figure 22, may be related with the position of the preceding sign. In Figure 23, the location of the target sign FARE in the third frame, which has the re-introduced discourse status, is assimilated to the location of the preceding (inflected) sign SEE-EACH.OTHER. However, the transition from the previous sign into the target sign (the frames between the second and third picture in Figure 23) is not considered to be part of the phonetic analysis, minimizing the effect of co-articulation in this case.

Figure 23. Co-articulation and position effect of the preceding sign on the target sign in re-introduction.

Finally, the monosyllabic nature of signs presents challenges in understanding phonological stress patterns, an area that remains understudied in TİD. Regarding internal phonological structure of the signs, in TİD, we know that, for example, reduplication serves various grammatical functions: in verbal signs, it is commonly used for verbal inflection (Karabuklu, 2022; Kelepir, 2020), while in nominal signs, it can indicate plurality. Although plural marking through reduplication was not required in the task design of this study, the possibility of other sign-internal phonological processes affecting duration cannot be excluded. For example, I noticed individual variation in the number of movement strokes for certain REs (e.g., some signers producing KEDI with one stroke while others used two or three strokes) and this phonological variation might warrant further investigation. Future research could examine the contribution of these factors to phonetic reduction patterns more systematically, particularly in relation to information structure and referential accessibility, while also considering the complex interplay between prosodic, phonological, and syntactic factors in signed narrative production.

Moving to the second research question, I had hypothesized that age of exposure to sign language might have some effects on the phonetic measures in an utterance-level analysis of narratives as a plethora of research has indicated differences among native and late signers on various components of language, including morphosyntax (Boudreault & Mayberry, 2006; Cormier, Schembri, Vinson, & Orfanidou, 2012; Newport, 1990), vocabulary development (Anderson, 2005; Cuetos, Monsalve, Pinto, & Rodríguez-Ferreiro, 2004; Jones et al., 2020; Keleş, Atmaca, & Gökgöz, 2022; Ramírez, Lieberman, & Mayberry, 2013), sentence comprehension (Cheng & Mayberry, 2019, 2021; A. M. Lieberman, Borovsky, & Mayberry, 2018), spatial development (Karadöller, Sümer, & Özyürek, 2020, 2023), and narrative competence (Becker, 2009; Cormier et al., 2013; Gür, 2024; Keleş et al., 2023).

The results of the utterance-level analysis revealed that the only significant phonetic measure was duration. Natives were significantly faster in the narration of the same event than late signers in TİD. There was no difference in the other phonetic measures, hand distance, and sign space use among the acquisition groups. There might be different factors contributing to the observed difference in the duration of event retellings. The second phonetic measure reported in this thesis, the accumulated hand distance, was calculated as the Euclidean distance of the hand coordinates between each frame, and it was normalized by the duration. Since time is controlled with this normalization, this measure (i.e., distance divided by time) can correspond to the average speed of the hand in event retellings. Since I did not observe a significant difference between the groups in the signing speed, but late signers had longer retelling durations than natives, this may imply that native signers have shorter durations not because they sign faster but perhaps due to other factors, such as more economical use of language (e.g., avoiding over-redundancy in reference tracking), as supported by the findings presented in Chapter 4.

To my knowledge, this is one of the first analyses that directly compares the phonetics of native and late signers using computer vision and pose estimation. For this reason, this thesis proposed a novel methodology for investigating the age of acquisition effects on phonetic parameters.

Given the paucity of relevant work, there are very few studies that compared the utterances of native and late signers with a non-computational phonetic analysis, with the exception of Braem (1999), who analyzed the personal narratives produced by early and late learners of Swiss German Sign Language (DSGS). In her study, late learners had slower speeds and smaller amplitude of the torso compared to natives. Yet, it is important to note that Braem did not include any computational methods in the analysis of the rhythmic patterns. In a previous study, we used pose estimation and found a significant difference in the amount of movement or articulatory energy for the non-dominant or left side of the body among native and late learners in complex motion events (Gökgöz & Keleş, 2023). In Flaherty et al. (2023), the NSL signers' year of entry to school was found to be significant in predicting the reduction in the signing space. Since younger NSL signers exhibited a more restricted use of signing space compared to older signers, the authors were able to document ongoing articulatory changes in the evolution of NSL. On a similar ground, Börstell et al. (2024) reported the effects of age and frequency on the average individual sign duration and signing rate of utterances for Swedish Sign Language, BSL, and NGT. In their work, older signers were slower in general in both measures than younger signers. The authors also investigated the effects of age of acquisition and parental status (i.e., whether the signers had deaf or hearing parents), both of which had no effect on sign duration and signing rate of utterances.

Considering this body of work, I had not predicted a reduction effect in the signing space among native and late signers in my study, as TİD is considered an established language. However, our results were in line with Braem's observations in DSGS because native TİD signers retold the same event more efficiently, which was reflected in a shorter overall duration.

The interpretation of this finding, nevertheless, comes with some caveats. Individual variability, such as differences in memory or subjective preferences in narrative planning and discourse coherence, might influence the observed variance in duration. For instance, some signers may have attended to details in one particular event of a narrative, retold more subcomponents, and even tracked referents in the background, whereas others might have simply focused on the protagonists and main components of the event. The latter preference may have resulted in a shorter retelling. To control for this, I classified each identified event retelling into three depending on how well it was remembered by the participants. Then, I selected the events that were consistently mentioned across all participants to ensure a relatively fair comparison in an attempt to minimize the influence of varying levels of detail in the retellings. In addition, while Börstell et al. (2024) demonstrated that a signer's age can significantly influence the duration of a sign and signing rate, it is important to note that the TİD signers in this study were age-matched. Also, even though Börstell et al. used a measure of an utterance's signing rate, this was defined as the number of signs divided by the duration of the utterance. As a result, they did not report any statistics on the duration of an utterance alone, which may explain why we obtained significant differences for this measure in our analysis. However, while this approach might mitigate some of these inconsistencies, the utterance-level phonetic analysis I present in this thesis should be interpreted with an understanding of these limitations. Therefore, more controlled experiments are needed to further understand the effects of delayed first language acquisition on articulatory phonetics.

Going back to the insignificant effect of hand distance and sign space use, we had previously seen in the phonetic analysis of complex motion events in TİD that being a late acquirer significantly increased the articulatory energy use on the non-dominant side (i.e., the left side of the body) (Gökgöz & Keleş, 2023). Following Sanders and Napoli's (2016) observations, we concluded that while native signers might not use the non-dominant side when it is not required, possibly due to an articulatory inhibition strategy, late signers less frequently employ this strategy. Nonetheless, like the duration calculation in Börstell et al. (2024), we had only investigated individual signs and did not report any measures on utterances in our pilot study. This contrastive finding may highlight the need for further investigation that includes both lexical-item-level and utterance-level metrics to fully understand the articulatory patterns of native and late signers in extended discourses.

CHAPTER 7

CONCLUSION

This thesis has explored discourse cohesion and phonetic reduction in Turkish Sign Language (TİD) narratives in an attempt to contribute to our understanding of language efficiency and referential accessibility. I used a quantized version of the theoretical framework of Ariel's (1990) accessibility hierarchy, following Czubek (2017) and Toole (1996), and investigated the discourse cohesion strategies of native and late signers in TİD narratives with a production experiment. This thesis also examined articulatory phonetic or kinematic patterns of referring expressions (REs) and utterances within narratives by employing MediaPipe, a markerless pose estimation library.

The thesis employed a production experiment, and the analyses on discourse cohesion and phonetics showed that TİD narrative data patterned well with the existing literature on cohesion strategies in other signed and spoken languages, particularly concerning the relationship between form and gradient referential accessibility. Overall, there was a tendency to choose reduced or more attenuated forms for the tracking of highly accessible referents and used more explicit forms like nominals and tracing tools to introduce or re-introduce referents with low accessibility. I have thus proposed a revised accessibility hierarchy of 8 different REs in TİD, with more explicit forms like nouns used for referents with low accessibility and more reduced forms like verbs for referents with high accessibility.

Age of acquisition significantly affected accessibility scores. Although the REs were used in a similar manner between the late and native signers, native acquisition led to increased accessibility in the retellings. In this context, a higher mean of accessibility indicated that the native signers usually maintained the same entity for a period of time, with fewer shifts to other competing referents. This might be due to a possible effect of age of acquisition on pragmatics, which could be explained with Sorace's (2006) syntax-discourse interface hypothesis, which was originally argued to be vulnerable, especially for second language speakers. In our case, this interface vulnerability may lead late signers to have difficulty efficiently selecting the appropriate referring expressions, resulting in over-explicitness or less efficient reference tracking in discourse, as evidenced by tracking referents with a lower mean of accessibility due to high engagement in topic shifts and more referential competition.

The thesis also included a computational phonetic analysis of the narratives using MediaPipe. The findings revealed that the phonetic measures, duration, hand distance, and sign space use were influenced by the discourse context and the gradient accessibility of referents, supporting the previous observations of phonetic reduction and linguistic efficiency in the literature for both spoken and sign languages. Specifically, more predictable or accessible referents displayed phonetic reduction, which surfaced as having less duration, less hand distance or movement, and a narrower sign space. However, our graded accessibility analysis allowed us to discover fine-grained observations in the phonetic parameters of TİD and how they might relate to pragmatics. With this analysis, we found that while kinematic measures were sensitive to a binary distinction between new versus predictable, duration was significant in all three discourse contexts, including re-introduction. Also, the graded accessibility analysis alluded to the possibility of rhythmic prosody in lists in TİD. The last item (e.g., referent) in three subsequent introductions underwent duration reduction. A further utterance-level phonetic examination of both acquisition groups indicated that late signers produced longer retellings than native signers as a response to the same stimuli. However, the age of acquisition was not a strong predictor of phonetic reduction since it was not uniformly observed across all kinematic measures. For instance, hand distance and sign space use did not show significant differences between native and late signers. This might mean that first language delay only affects linguistic efficiency surfacing in the duration parameter, and it does not create any difference in the amount of movement or signing space size. The kinematic measures can perhaps be an important predictor of late acquisition in the contexts of second language acquisition or cohorts acquiring a newly emerging language, such as NSL (Flaherty, Sato, & Kirby, 2023). More research is needed to justify this assumption.

Several limitations must be acknowledged. Firstly, the production experiment presented in this thesis was not interactive, as there was no interlocutor. The majority of the signers narrated by looking at the camera, but some preferred to narrate the clips to the research assistant, who did not actively respond or engage in. Although the effect of looking at the camera versus an addressee was insignificant (see Appendix B), we know that interaction theories (e.g., Gumpers, 2015) suggest that real-time feedback from an interlocutor could influence language production. Future research may address this issue by investigating whether interactional contexts (versus monologues) alter any of the discourse cohesion or phonetic reduction patterns observed in this study.

As another limitation, this thesis only investigated one narrative type, which was cartoon retellings, leaving other types such as personal experience narratives unexplored. While previous work found similar reference tracking strategies across different narrative types (Czubek, 2017), future research can look at how TİD signers produce other types of narratives.

In addition, the concurrent forms, or simultaneous tracking of multiple entities, were beyond the scope of this study since I have only annotated the subject referents for maintenance and re-introduction. Most often, however, sign languages can encode the character perspective (e.g., use constructed action and track the agent on an action) and the observer perspective (e.g., use a manual classifier sign to track the patient referent) simultaneously by capitalizing on the affordances of the visual-manual modality, and these two perspectives can blend together (Dudis, 2004; Janzen, 2004; Smith & Cormier, 2014). Similarly, I also did not consider the role of animacy of referents, both of which might affect the cohesion strategies of the signers (Ferrara et al., 2022; Hodge, Ferrara, & Anible, 2019). It is possible to investigate both of these factors in future work.

To my knowledge, this thesis reported the first computational phonetic analysis of TİD narratives. By providing a systematic, data-driven approach to analyzing the phonetic properties of TİD, we aimed to provide insights into the relationship between phonetic reduction, referential accessibility, and discourse cohesion. But I must acknowledge certain limitations. In the computational analysis, the continuous phonetic measures like duration, hand distance, and sign space use were directly mapped to a cognitive variable (i.e., referential accessibility) without an intermediary phonological grouping, possibly resulting in the mixing of phonological and phonetic factors. This could limit our interpretation of the results since phonetic variations can be affected by phonological or prosodic components that were not taken into consideration in the analysis. For example, in spoken languages, changes in duration are often accompanied by other correlates of prosody (such as changes in amplitude and fundamental frequency) (Baker & Bradlow, 2009; Edwards & Beckman, 1988) and can be due to lexical (e.g., stress) or sentence level (e.g., information structure such as being focused versus non-focused) prosody (Breen, Fedorenko, Wagner, & Gibson, 2010). Similarly, TİD signers may be affected by these cues, including pitch variations (as might be indicated through non-manual markers) or distinctions between being focused and non-focused. Although there are certain phonological and prosodic models proposed for sign languages (Brentari, 1998; Van der Hulst, 1993), there is no consensus on the extent of such phonological and prosodic groupings. It is also difficult to determine the information status of a referent (focus versus topic) only based on prosody or non-manual markers. For instance, prosodic marking on a topic seems to be optional in TİD (Keleş & Gökgöz, 2022). Nonetheless, it is important for future research to take into account this limitation when analyzing continuous phonetic data for sign languages.

APPENDIX A

INFORMED CONSENT FORM

KATILIMCI BİLGİ ve ONAM FORMU

Boğaziçi Üniversitesi

Dilbilimsel Araştırmalara Dayanan Okul Öncesi Hikayeler İle Anne-babası Duyup Kendisi Sağır Olan Çocukların İşaret Dili Gelişimini Desteklemek

İngilizce adı

Supporting Sign Language Development of Deaf Children with Hearing Parents through Linguistically-Informed Preschool Stories

KADIR GÖKGÖZ

Adresi: Boğaziçi Üniversitesi, Fen-Edebiyat Fakültesi, Dilbilim Bölümü, John Freely Binası, Oda No 306, 34342, Bebek/İstanbul, Türkiye

E-mail adresi: kadir.gokgoz@boun.edu.tr

Telefonu: 0212359-4902

Sayın katılımcı,

Bu projenin birinci konusu Türk İşaret Dili'ni anlamak ve İşaret Dilini kullanan bireylerin dil kullanımları arasında ne gibi farklılıklar var bunları görmek.

İkinci konusu, birinci konuda öğrendiklerimizi kullanmak ve sağır çocuklar için İşaret Dili Hikayeleri oluşturmak.

Siz projenin birinci konusunda yer alacaksınız.

Bu projeyi yürüten Boğaziçi Üniversitesi Dilbilim Bölümü araştırmacıları. Araştırma yürütücüsünün adı Kadir Gökgöz. Türk İşaret Dili dilbilgisi kurallarını araştıracağız. Bunu yaparken sizinle bazı çalışmalar yapacağız. Örnekler:

1. resme bakarak resmi tarif etmek, resimde neler oluyor anlatmak

2. video seyretmek ve orada neler oluyor anlatmak

Bunlar dışında başka bir işitme engelli katılımcı ile bir konuda konuşmak olabilir.

Bazı çalışmalarda size Türk İşaret Dili'nde bir cümlenin filmini gösterip "Bu cümle doğru mu?", "Siz söyler misiniz?" gibi sorular sorabiliriz.

Siz cevap verirken sizi videoya çekeceğiz. Bu videolar bizim için araştırma verisi oluşturmaktadır.

Bu görüntü verilerini kilitli ve şifreli depolarda koruyacağız ve kimse projedeki araştırmacıların izni olmadan bu videoları kullanamayacak.

Veriler sadece araştırma amaçlı kullanılacaktır. Siz izin verirseniz video verileri veya fotoğraf görüntü verileriniz bilimsel amaçlı olarak incelenebilir ve bilimsel yayın ve konferanslarda gösterilebilir.

Siz izin verirseniz bu görüntüler projenin internet sitesinde gösterilebilir. Bu siteye herkes bakabilir.

1. Projeye katılım sizin isteğinize bağlıdır.

2. Çalışmalar 1-2 saat sürebilir. Uzun çalışmalarda sık sık ara vereceğiz. Siz istediğiniz zaman mola verebilirsiniz. Çalışmayı istediğiniz zaman bitirebilirsiniz.

3. Söylediğiniz bir şey kayıttan silinsin isterseniz lütfen bize söyleyin, o bölümü çıkarırız.

4. Çalışmadan çekilmeye karar verirseniz, video ve fotoğraf verileriniz tamamen imha edilecektir. Bugünden sonra başka bir gün video görüntüleriniz ile ilgili fikrinizi değiştirirseniz ve kullanmamızı istemezseniz bize haber verin, bu görüntüleri kullanmayacağız. Bu veriler de imha edilecektir.

5. Siz istemezseniz isminiz ve kimliğiniz hiçbir yerde yayınlanmayacak.

6. Katılım sonucu, zamanınızın karşılığı olarak size ödeme yapılacaktır. Bu ödeme banka hesabınıza transfer şeklinde olacaktır.

Sayın katılımcı,

Yukarıdaki bilgileri veya size gösterilen videoda anlatılanları anladınız mı?
(Evet / Hayır)

Eğer projemize katılmayı kabul ediyorsanız lütfen aşağıdaki sorulara cevap verin ve daha sonra bu formu imzalayın.

Bu imzalı formdan bir tane sizde bir tane bizde kalacak.

(a) Bu araştırma için video görüntülerinizin çekilmesine izin veriyor musunuz?
(Evet / Hayır)

(b) Video görüntülerinizin araştırma amacıyla projenin internet sayfasında herkese açık olarak gösterilmesine izin veriyor musunuz?

(Evet / Hayır)

(c) Video görüntülerinizin bilimsel yayınlarda ve konferanslarda kullanılmasına izin veriyor musunuz?

(Evet / Hayır)

(d) Şunu kabul ediyor musunuz?

“Video görüntülerim ücretsiz olarak görülebilir. Kimse bu video görüntülerinden para kazanamaz. Ancak bu görüntüler kullanılarak yeni şeyler yapılabilir. Örneğin, görüntüler internet sitelerinde veya konferans sunumlarında kullanılabilir. Bu ürünler de para kazanmak için kullanılamaz.”

(Evet / Hayır)

(e) Video görüntülerinizin sonsuza dek kullanılmasına (internette veya başka bir araçla) izin veriyor musunuz?

(Evet / Hayır)

(f) İsminizi bilimsel yayınlardaki ve projenin internet sitesindeki "Teşekkür" bölümüne yazmamıza izin veriyor musunuz?

(Evet / Hayır)

(g) Video görüntülerinizin araştırma amacıyla Türkiye ve başka ülkelerdeki araştırmacılar ile paylaşılmasına izin veriyor musunuz? Kabul ederseniz video görüntüleriniz bu araştırmacı ve ülkelere gidecek.

(Evet / Hayır)

“Formu okudum. Bana anlatılanları anladım. Bu projeye kendi isteğimle katılıyorum. Beni kimse zorlamıyor. Bu formun bir kopyasını araştırmacıdan aldım.”

Adınız, Soyadınız	İmzanız	Tarih
....	....	....

APPENDIX B

LOOKING AT THE CAMERA VERSUS ADDRESSEE MODELS

Table B.1. Discourse Model: Model results for Accessibility Score

Predictors	Estimates	CI	p
(Intercept)	1.49	1.37 – 1.62	<0.001
Group [Native]	0.08	0.03 – 0.13	0.001
Discourse [Introduction]	-2.28	-2.31 – -2.24	<0.001
Discourse [Maintenance]	2.68	2.65 – 2.71	<0.001
Interaction [Camera]	-0.02	-0.09 – 0.05	0.514
Random Effects			
σ^2	0.60		
τ_{00} Subject	0.01		
τ_{00} Narrative	0.03		
ICC	0.06		
N Subject	28		
N Narrative	10		
Observations	4027		
Marginal R^2 / Conditional R^2	0.869 / 0.877		

Table B.2. RE Type Model: Model results for Accessibility Score

Predictors	Estimates	CI	p
(Intercept)	2.52	2.30 – 2.73	<0.001
Group [Native]	0.16	0.08 – 0.24	<0.001
RE Type [NOM]	-2.29	-2.36 – -2.22	<0.001
RE Type [CL]	0.33	0.21 – 0.44	<0.001
RE Type [CA]	0.75	0.66 – 0.84	<0.001
Interaction [Camera]	-0.09	-0.19 – -0.02	0.120
Group [Native] × RE Type [NOM]	-0.11	-0.18 – -0.04	0.004
Group [Native] × RE Type [CL]	0.10	-0.01 – 0.21	0.075
Group [Native] × RE Type [CA]	0.01	-0.08 – 0.10	0.787
Random Effects			
σ^2	2.35		
τ_{00} Subject	0.02		
τ_{00} Narrative	0.09		
ICC	0.05		
N Subject	28		
N Narrative	10		
Observations	4027		
Marginal R^2 / Conditional R^2	0.499 / 0.521		

APPENDIX C

OUTLIER ANALYSIS

Prior to the analysis of the accessibility values and phonetic measures in the computer vision analysis, I performed an outlier analysis to the count data to see whether this signer significantly diverged from other signers in terms of the distribution of RE types. Figure C1 shows a comparison of the proportions of different RE types (discourse contexts collapsed) used by a potential outlier (participant numbered 34) and all the other signers. As can be seen, Participant 34 predominantly used CA, accounting for 78% of their total REs. This contrasts with the average CA use among other signers, which is only 24%. Similarly, Participant 34 only uses nominal constructions 6% of the time in the production experiment whereas nominals are the main REs among other participants.

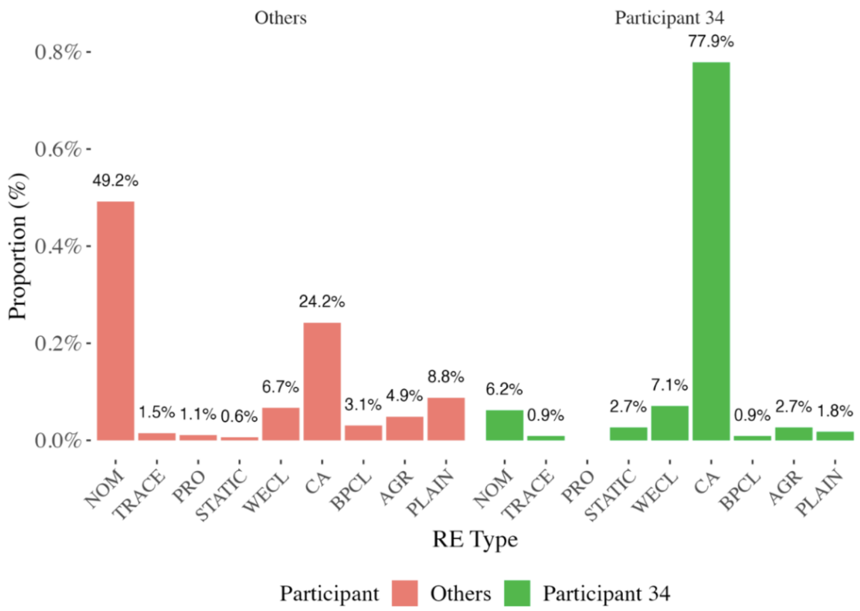


Figure C1. Proportions of RE Type used by Participant 34 and other participants (on average)

If this signer statistically differs from the rest of the participants, excluding them from the accessibility analysis will be necessary since the model predictions must reflect the typical behavior of the participant group without being disproportionately influenced by an atypical data point. To achieve this, I dichotomized RE Type into two levels: NOM and NON-NOM. I then fit a generalized linear mixed-effects model (GLMM) using the `glmer` function from the `lme4` (1.1.27) package to the binary RE Type (modeled using a binomial distribution) with Acquisition and Discourse as fixed effects, as well as random intercepts for Participant and Stimuli:

(1) Diagnostic Model = glmer(BinaryRefType   Acquisition + Discourse + (1|Participant) + (1|Stimuli))

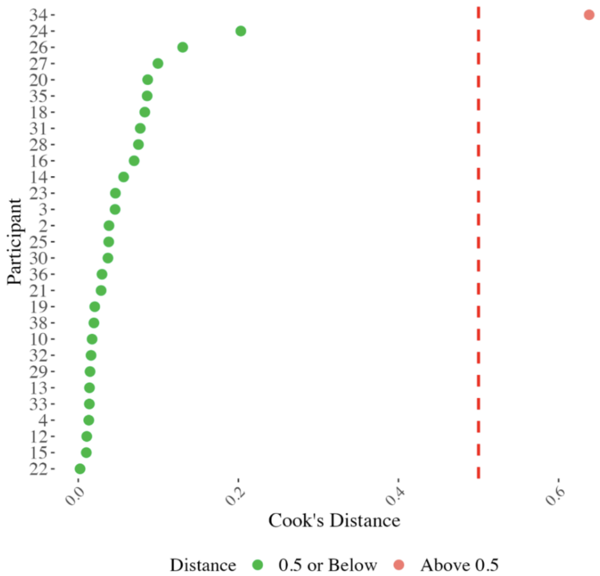


Figure C2. Cook's Distance by participant for outlier analysis

Then, I assessed the influence of individual participants on the model's estimates by calculating Cook's distance with the `influence.me` (0.9-9) package. This diagnostic measure is used to identify random factors that had a disproportionate impact on the model's parameters. In my case, I considered participants who received a distance score of 0.5, which is the conventional threshold in the literature, and greater to be influential on the model. Figure C2 illustrates the Cook's Distance values for each participant in the model, quantifying the influence on the nominal versus non-nominal use in the narrative production.

As indicated by the outlier analysis, Participant 34 had the greatest influence with a Cook's distance of approximately 0.6. As a result, I fit the same model this time by excluding this participant and compared both models (Tables C1 and C2).

Table C1. Model results with Participant 34 included

Predictors	Odds Ratios	CI	p
Intercept	1.04	0.73 – 1.48	0.822
Group [Native]	0.89	0.65 – 1.21	0.451
Discourse [Introduction]	14.63	12.07 – 17.73	<0.001
Discourse [Maintenance]	0.04	0.04 – 0.05	<0.001

Table C2. Model results with Participant 34 excluded

Predictors	Odds Ratios	CI	p
Intercept	1.20	0.92 – 1.55	0.175
Group [Native]	1.00	0.82 – 1.23	0.976
Discourse [Introduction]	14.65	12.04 – 17.83	<0.001
Discourse [Maintenance]	0.04	0.04 – 0.05	<0.001

APPENDIX D

MULTINOMIAL ANALYSIS

Table D1. Multinomial model results for RE Type

Predictors	Odds Ratios	CI	p	Response
Accessibility	1.19	0.99 – 1.44	0.070	NOM
Discourse [Introduction]	9.45	4.92 – 18.16	<0.001	NOM
Discourse [Maintenance]	0.03	0.01 – 0.10	<0.001	NOM
Group [Native]	1.09	0.97 – 1.23	0.143	NOM
Accessibility × Discourse [Introduction]	1.56	1.20 – 2.03	0.001	NOM
Accessibility × Discourse [Maintenance]	1.00	0.75 – 1.34	0.976	NOM
Accessibility	1.72	0.93 – 3.18	0.084	CA
Discourse [Introduction]	0.08	0.02 – 0.27	<0.001	CA
Discourse [Maintenance]	4.56	1.95 – 10.66	<0.001	CA
Group [Native]	1.16	1.03 – 1.29	0.010	CA
Accessibility × Discourse [Introduction]	3.66	1.09 – 12.24	0.035	CA
Accessibility × Discourse [Maintenance]	0.49	0.26 – 0.92	0.025	CA
Accessibility	1.72	1.06 – 2.78	0.029	VERB
Discourse [Introduction]	0.29	0.10 – 0.83	0.022	VERB
Discourse [Maintenance]	2.07	0.88 – 4.87	0.094	VERB
Group [Native]	1.22	1.08 – 1.39	0.002	VERB
Accessibility × Discourse [Introduction]	2.02	0.80 – 5.09	0.137	VERB
Accessibility × Discourse [Maintenance]	0.69	0.42 – 1.15	0.154	VERB
Observations	4027			
R^2 / R^2 adjusted	0.299 / 0.299			

REFERENCES

Açan, A. Z. (2007). A linguistic analysis on basic sentence types in turkish sign language (TİD) with reference to non-manual activity. *Unpublished doctorate thesis, University of Hacettepe, Ankara*.

Ahlgren, I., & Bergman, B. (1994). Reference in narratives. In I. Ahlgren, B. Bergman, & M. Brennan (Eds.), *Perspectives on sign language structure: Papers from the fifth international symposium on sign language research* (pp. 29–36). ISLA.

Ahrenholz, B. (2005). Reference to persons and objects in the function of subject in learner varieties [Publisher: Mouton de Gruyter Berlin, Germany]. *The structure of learner varieties*, 19–64.

Aksu-Koç, A., & Nicolopoulou, A. (2015). Character reference in young children's narratives: A crosslinguistic comparison of English, Greek, and Turkish. *Lingua*, *155*, 62–84. doi:10.1016/j.lingua.2014.04.006.

Anderson, D. (2005). *Lexical Development of Deaf Children Acquiring Signed Languages* [Publication Title: Advances in the Sign-Language Development of Deaf Children Section: Advances in the Sign-Language Development of Deaf Children]. Oxford University Press.

Ann, J. (2006). *Frequency of occurrence and ease of articulation of sign language handshapes: The Taiwanese example*. Gallaudet University Press. Retrieved October 13, 2024, from https://muse.jhu.edu/pub/18/monograph/book/5254

Ariel, M. (1990). Accessing Noun-Phrase Antecedents. Vol. 96. Londres: Routlege. *Linguistics*, 113–118.

Aylett, M., & Turk, A. (2004). The Smooth Signal Redundancy Hypothesis: A Functional Explanation for Relationships between Redundancy, Prosodic Prominence, and Duration in Spontaneous Speech. *Language and Speech*, *47*(1), 31–56. doi:10.1177/00238309040470010201.

Azar, Z., & Özyürek, A. (2015). Discourse management: Reference tracking in speech and gesture in Turkish narratives. *Dutch Journal of Applied Linguistics*, *4*(2), 222–240. doi:10.1075/dujal.4.2.06aza.

Bahan, B., Kegl, J., MacLaughlin, D., & Neidle, C. (1995). Convergent evidence for the structure of determiner phrases in American Sign Language. *FLSM VI: Proceedings of the Sixth Annual Meeting of the Formal Linguistics Society of Mid-America*, *2*, 1–12.

Baker, R. E., & Bradlow, A. R. (2009). Variability in Word Duration as a Function of Probability, Speech Style, and Prosody. *Language and Speech*, *52*(4), 391–413. doi:10.1177/0023830909336575.

Baltrusaitis, T., Zadeh, A., Lim, Y. C., & Morency, L.-P. (2018). OpenFace 2.0: Facial Behavior Analysis Toolkit. *2018 13th IEEE International Conference* *on Automatic Face & Gesture Recognition (FG 2018)*, 59–66. doi:10.1109/FG.2018.00019.

Barberà, G. (2013). The meaning of space in Catalan Sign Language (LSC): Reference, specificity and structure in signed discourse. *Sign Language & Linguistics*, *16*(1), 97–105. doi:10.1075/sll.16.1.05bar.

Barberà, G., & Massó, G. (2009). Keeping track in sign language: Coreference annotation using Centering Theory. *Talk presented at the Sign Language Corpora: Linguistic issues workshop, University College London*.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue [Publisher: Elsevier]. *Journal of memory and language*, *42*(1), 1–22.

Bates, D., Maechler, M., Bolker [aut, B., cre, Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., Bauer, A., simulate.formula), P. N. K. (c. o., Tanaka, E., & Jagan, M. (2024, July). Lme4: Linear Mixed-Effects Models using 'Eigen' and S4. Retrieved July 14, 2024, from https://cran.r-project.org/web/packages/lme4/index.html

Bauer, A., Kuder, A., Schulder, M., & Schepens, J. (2024). Phonetic differences between affirmative and feedback head nods in German Sign Language (DGS): A pose estimation study [Publisher: Public Library of Science]. *PLOS ONE*, *19*(5), e0304040. doi:10.1371/journal.pone.0304040.

Becker, C. (2009). Narrative competences of Deaf children in German Sign Language. *Sign Language & Linguistics*, *12*(2), 113–160. doi:10.1075/sll.12.2.02bec.

Bel, A., Ortells, M., & Morgan, G. (2015). Reference control in the narratives of adult sign language learners. *International Journal of Bilingualism*, *19*(5), 608–624. doi:10.1177/1367006914527186.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*(1), 92–111. doi:10.1016/j.jml.2008.06.003.

Berman, R. A., & Slobin, D. I. (2013). *Relating events in narrative: A crosslinguistic developmental study*. Psychology Press.

Börstell, C. (2023, May). Extracting Sign Language Articulation from Videos with MediaPipe. In T. Alumäe & M. Fishel (Eds.), *Proceedings of the 24th Nordic Conference on Computational Linguistics (NoDaLiDa)* (pp. 169–178). University of Tartu Library. https://aclanthology.org/2023.nodalida-1.18

Börstell, C., & Lepic, R. (2020). Spatial metaphors in antonym pairs across sign languages. *Sign Language & Linguistics*, *23*(1-2), 112–141. doi:10.1075/sll.00046.bor.

Börstell, C., Schembri, A., & Crasborn, O. (2024). Sign duration and signing rate in British Sign Language, Dutch Sign Language and Swedish Sign Language. *Glossa Psycholinguistics*, *3*(1). doi:10.5070/G60111915.

Boudreault, P., & Mayberry, R. I. (2006). Grammatical processing in American Sign Language: Age of first-language acquisition effects in relation to syntactic structure. *Language and Cognitive Processes*, *21*(5), 608–635. doi:10.1080/01690960500139363.

Braem, P. B. (1999). Rhythmic Temporal Patterns in the Signing of Deaf Early and Late Learners of Swiss German Sign Language. *Language and Speech*, *42*(2-3), 177–208. doi:10.1177/00238309990420020301.

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, *25*(7-9), 1044–1098. doi:10.1080/01690965.2010.504378.

Brehm, L., & Alday, P. M. (2022). Contrast coding choices in a decade of mixed models. *Journal of Memory and Language*, *125*, 104334.

Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology*. MIT Press.

Brozdowski, C., & Emmorey, K. (2023). Using transitional information in sign and gesture perception. *Acta psychologica*, *236*, 103923.

Bürkner, P.-C. (2019). Bayesian item response modeling in r with brms and stan. *arXiv preprint arXiv:1905.09501*.

Bybee, J. L. (2003). *Phonology and language use*. Cambridge University Press.

Calderone, C. (2020). *Can you retrieve it? : Pragmatic, morpho-syntactic and prosodic features in sentence topic types in italian sign language (lis)* [Doctoral Thesis]. Università Ca' Foscari Venezia.

Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y. (2017). Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields, 7291–7299.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. *Subject and Topic*.

Cheng, Q., & Mayberry, R. I. (2019). Acquiring a first language in adolescence: The case of basic word order in American Sign Language. *Journal of Child Language*, *46*(2), 214–240. doi:10.1017/S0305000918000417.

Cheng, Q., & Mayberry, R. I. (2021). When event knowledge overrides word order in sentence comprehension: Learning a first language after childhood. *Developmental Science*. doi:10.1111/desc.13073.

Cogill-Koez, D. (2000). Signed language classifier predicates: Linguistic structures or schematic visual representation? *Sign Language & Linguistics*, *3*(2), 153–207. doi:10.1075/sll.3.2.03cog.

Colyer, S. L., Evans, M., Cosker, D. P., & Salo, A. I. T. (2018). A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System. *Sports Medicine - Open*, *4*(1), 24. doi:10.1186/s40798-018-0139-y.

Cormier, K., Fenlon, J., & Schembri, A. (2015). Indicating verbs in British Sign Language favour motivated use of space. *Open Linguistics*, *1*(1). doi:10.1515/opli-2015-0025.

Cormier, K., Schembri, A., Vinson, D., & Orfanidou, E. (2012). First language acquisition differs from second language acquisition in prelingually deaf signers: Evidence from sensitivity to grammaticality judgement in British Sign Language. *Cognition*, *124*(1), 50–65. doi:10.1016/j.cognition.2012.04.003.

Cormier, K., Smith, S., & Zwets, M. (2013). Framing constructed action in British Sign Language narratives. *Journal of Pragmatics*, *55*, 119–139. doi:10.1016/j.pragma.2013.06.002.

Crasborn, O. (2012). Phonetics. In *Sign language: An international handbook*. De Gruyter.

Crasborn, O., Van der Kooij, E., & Ros, J. (2012). On the weight of phrase-final prosodic words in a sign language. *Sign Language & Linguistics*, *15*(1), 11–38.

Cuetos, F., Monsalve, A., Pinto, A., & Rodríguez-Ferreiro, J. (2004). Predictor variables of written picture naming in the deaf. *Reading and Writing*, *17*(3), 227–240. Retrieved April 11, 2020, from insights.ovid.com

Czubek, T. A. (2017). *A comprehensive study of referring expressions in asl* [Doctoral dissertation, Boston University].

Daniloff, R. G., & Hammarberg, R. E. (1973). On defining coarticulation. *Journal of Phonetics*, *1*(3), 239–248. doi:10.1016/S0095-4470(19)31388-9.

Debreslioska, S., Özyürek, A., Gullberg, M., & Perniss, P. (2013). Gestural Viewpoint Signals Referent Accessibility [Publisher: Routledge _eprint: https://doi.org/10.1080/0163853X.2013.824286]. *Discourse Processes*, *50*(7), 431–456. doi:10.1080/0163853X.2013.824286.

Dikyuva, H., Makaroğlu, B., & Arık, E. (2015). Türk işaret dili dilbilgisi kitabı. *Aile ve Sosyal Politikalar Bakanlığı Yayınları: Ankara*.

Dudis, P. G. (2004). Body partitioning and real-space blends. *Cognitive Linguistics*, *15*(2). doi:10.1515/cogl.2004.009.

Edwards, J., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *the Journal of the Acoustical Society of America*, *89*(1), 369–382.

Edwards, J., & Beckman, M. (1988). Articulatory Timing and the Prosodic Interpretation of Syllable Duration. *Phonetica*, *45*(2-4), 156–174. doi:10.1159/000261824.

Emmorey, K. (2003). *Perspectives on classifier constructions in sign languages*. Psychology Press.

Emmorey, K., Norman, F., & O'grady, L. (1991). The activation of spatial antecedents from overt pronouns in american sign language. *Language and Cognitive Processes*, *6*(3), 207–228. doi:10.1080/01690969108406943.

Engberg-Pedersen, E. (1993). Space in Danish Sign Language: The semantics and morphosyntax of the use of space in a visual language [Publisher: SIGNUM-Press].

Erickson, F. (1992). They know all the lines: Rhythmic organization and contextualization in a conversational listing routine. *The contextualization of language*, 365–397.

Eroğlu, E., Şahin, K., Kisbet, B., & Gökgöz, K. (2024). Pointing in Turkish Sign Language (TİD) [Poster presented at the IMPRS Conference, Nijmegen].

Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y., & Tu, X. M. (2014). Log-transformation and its implications for data analysis. *Shanghai archives of psychiatry*, *26*(2), 105–109. Retrieved September 25, 2024, from https://europepmc.org/articles/4120293

Ferrara, L., Anible, B., Hodge, G., Jantunen, T., Leeson, L., Mesch, J., & Nilsson, A.-L. (2022). A cross-linguistic comparison of reference across five signed languages [Publisher: De Gruyter Mouton]. *Linguistic Typology*.

Ferrero, G. (1894). L'inertie mentale et la loi du moindre effort [Publisher: JSTOR]. *Revue Philosophique de la France et de l'Étranger*, *37*, 169–182.

Flaherty, M., Sato, A., & Kirby, S. (2023). Documenting a Reduction in Signing Space in Nicaraguan Sign Language Using Depth and Motion Capture. *Cognitive Science*, *47*(4), e13277. doi:10.1111/cogs.13277.

Fowler, C. A. (1988). Differential Shortening of Repeated Content Words Produced in Various Communicative Contexts. *Language and Speech*, *31*(4), 307–319. doi:10.1177/002383098803100401.

Fox, J., Friendly, M., & Weisberg, S. (2013). Hypothesis tests for multivariate linear models using the car package. *The R Journal*, *5*(1), 39.

Fragkiadakis, M. (2022). Assessing an Automated Tool to Quantify Variation in Movement and Location [Publisher: JSTOR]. *Sign Language Studies*, *23*(1), 98–126.

Frederiksen, A. T., & Mayberry, R. I. (2016). Who's on First? Investigating the referential hierarchy in simple native ASL narratives. *Lingua*, *180*, 49–68. doi:10.1016/j.lingua.2016.03.007.

Frederiksen, A. T., & Mayberry, R. I. (2019). Reference tracking in early stages of different modality L2 acquisition: Limited over-explicitness in novice ASL signers' referring expressions. *Second Language Research*, *35*(2), 253–283. doi:10.1177/0267658317750220.

Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language*, *66*(4), 789–806. doi:10.1016/j.jml.2011.11.006.

Garcia, B., & Sallandre, M.-A. (2014). Reference resolution in French Sign Language. In P. C. H.
bibinitperiod A. Zribi-Hertz (Ed.), *Crosslinguistic studies on Noun Phrase structure and reference* (pp. 316–364). Brill.

Gerwing, J., & Bavelas, J. (2005). Linguistic influences on gesture's form. *Gesture*, *4*(2), 157–195. doi:10.1075/gest.4.2.04ger.

Givon, T. (1983). *Topic Continuity in Discourse*. John Benjamins Publishing Company. Retrieved September 20, 2020, from https://benjamins.com/catalog/tsl.3

Gökgöz, K. (2009). *Topics in turkish sign language (türk işaret dili–tid) syntax: Verb movement, negation and clausal architecture* [Master's thesis, Boğaziçi University].

Gökgöz, K. (2013). Negation in turkish sign language: The syntax of nonmanual markers. *Nonmanuals in Sign Language. Philadelphia, Benjamins*, 47–72.

Gökgöz, K. (2024). Verbal classifiers from a crosslinguistic and cross-modal point of view. *Language*, *100*(2), 179–217.

Gökgöz, K., & Keleş, O. (2023). Morpho-phonology and articulatory energy in expressing complex motion events in turkish sign language (TİD) and age of acquisition effects [Poster presented at the Boston University Conference on Language Development (BUCLD) Conference, Boston].

Göksel, A., Kelepir, M., Pfau, R., Steinbach, M., & Herrmann, A. (2016). Observations on clausal complementation in turkish sign language. *A matter of complexity: Subordination in sign languages*, 65–94.

Goldin-Meadow, S. (2005). *The Resilience of Language: What Gesture Creation in Deaf Children Can Tell Us about how All Children Learn Language*. Psychology Press.

Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41–58). Brill.

Grosvald, M. (2009). *Long-distance coarticulation: A production and perception study of English and American Sign Language*. University of California, Davis. Retrieved October 13, 2024, from https://search.proquest.com/openview/ba523e865baedcbfae3d5a10c0a645fb/1?pq-origsite=gscholar&cbl=18750

Grosvald, M., & Corina, D. P. (2012). Exploring the movement dynamics of manual and oral articulation: Evidence from coarticulation [Publisher: De Gruyter Mouton]. *Laboratory Phonology*, *3*(1), 37–60. doi:10.1515/lp-2012-0004.

Grosz, B., Joshi, A., & Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse [Publisher: Association for Computational Linguistics/MIT]. *Computational linguistics*. Retrieved October 6, 2024, from https://dash.harvard.edu/handle/1/2562369

Gullberg, M. (2006). Handling Discourse: Gestures, Reference Tracking, and Communication Strategies in Early L2. *Language Learning*, *56*(1), 155–196. doi:10.1111/j.0023-8333.2006.00344.x.

Gumperz, J. J. (2015). Interactional sociolinguistics a personal perspective. *The handbook of discourse analysis*, 309–323.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse [Publisher: Linguistic Society of America]. *Language*, *69*(2), 274–307. doi:10.2307/416535.

Gür, C. (2024). Investigating the effects of late sign language acquisition on referent introduction: A follow-up study. *Poznan Studies in Contemporary Linguistics*, *60*(1), 1–25. doi:10.1515/psicl-2022-1024.

Gür, C., & Sümer, B. (2022). Learning to introduce referents in narration is resilient to the effects of late sign language exposure. *Sign Language & Linguistics*, *25*(2), 205–234. doi:10.1075/sll.21004.gur.

Hendriks, H. (2003). Using nouns for reference maintenance: A seeming contradiction in l2 discourse. *Typology and second language acquisition*, 291–326.

Hickmann, M., & Hendriks, H. (1999). Cohesion and anaphora in children's narratives: A comparison of English, French, German, and Mandarin Chinese. *Journal of Child Language*, *26*(2), 419–452. doi:10.1017/S0305000999003785.

Hickmann, M., Hendriks, H., Roland, F., & Liang, J. (1996). The marking of new information in children's narratives: A comparison of English, French, German and Mandarin Chinese [Publisher: Cambridge University Press]. *Journal of Child Language*, *23*(3), 591–619. doi:10.1017/S0305000900008965.

Hodge, G., Ferrara, L. N., & Anible, B. D. (2019). The semiotic diversity of doing reference in a deaf signed language [Publisher: Elsevier]. *Journal of Pragmatics*, *143*, 33–53.

Hoetjes, M., Krahmer, E., & Swerts, M. (2014). Do repeated references result in sign reduction? *Sign Language & Linguistics*, *17*(1), 56–81. doi:10.1075/sll.17.1.03hoe.

Horn, L. (1986). Presupposition, theme and variations. *Chicago Linguistic Society*, *22*, 168–192.

Ilkbasaran, D. (2015). *Literacies, Mobilities and Agencies of Deaf Youth in Turkey: Constraints and Opportunities in the 21st Century* [Ph.D.] [ISBN: 9781321548945].

Jantunen, T., Burger, B., De Weerdt, D., Seilola, I., & Wainio, T. (2012). Experiences collecting motion capture data on continuous signing. *Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon*, 75–82.

Jantunen, T., De Weerdt, D., Burger, B., & Puupponen, A. (2020). The more you move, the more action you construct: A motion capture study on head and upper-torso movements in constructed action in Finnish Sign Language narratives. *Gesture*, *19*(1), 72–96. doi:10.1075/gest.19042.jan.

Jantunen, T., Puupponen, A., & Burger, B. (2020, May). What Comes First: Combining Motion Capture and Eye Tracking Data to Study the Order of Articulators in Constructed Action in Sign Language Narratives. In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 6003–6007). European Language Resources Association. Retrieved October 10, 2024, from https://aclanthology.org/2020.lrec-1.735

Janzen, T. (2004). Space rotation, perspective shift, and verb morphology in ASL. *Cognitive Linguistics*, *15*(2). doi:10.1515/cogl.2004.006.

Jedlička, P., Krňoul, Z., Kanis, J., & Železný, M. (2020, May). Sign Language Motion Capture Dataset for Data-driven Synthesis. In E. Efthimiou, S.-E. Fotinea, T. Hanke, J. A. Hochgesang, J. Kristoffersen, & J. Mesch (Eds.), *Proceedings of the LREC2020 9th Workshop on the Representation and Processing of Sign Languages: Sign Language Resources in the Service of the Language Community, Technological Challenges and Application Perspectives* (pp. 101–106). European Language Resources Association (ELRA). Retrieved October 14, 2024, from https://aclanthology.org/2020.signlang-1.16

Johnson-Laird, P. N. (1988). *The Computer and the Mind: An Introduction to Cognitive Science*. Harvard University Press.

Jones, A., Atkinson, J., Marshall, C., Botting, N., Clair, M. C. S., & Morgan, G. (2020). Expressive Vocabulary Predicts Nonverbal Executive Function: A 2-year Longitudinal Study of Deaf and Hearing Children [_eprint: https://srcd.onlinelibrary.wiley.com/doi/pdf/10.1111/cdev.13226]. *Child Development*, *91*(2), e400–e414. doi:10.1111/cdev.13226.

Jurafsky, D. (2003). Probabilistic modeling in psycholinguistics: Linguistic comprehension and production [Probabilistic Linguistics].

Karabuklu, S. (2022). *Modal signs and cooccurring nonmanual markers in turkish sign language (tid)* [Doctoral dissertation, Purdue University].

Karadöller, D. Z., Sümer, B., & Özyürek, A. (2020). Effects and Non-Effects of Late Language Exposure on Spatial Language Development: Evidence from Deaf Adults and Children [Publisher: Routledge _eprint: https://doi.org/10.1080/15475441.2020.1823846]. *Language Learning and Development*, *0*(0), 1–25. doi:10.1080/15475441.2020.1823846.

Karadöller, D. Z., Sümer, B., Ünal, E., & Özyürek, A. (2023). Late sign language exposure does not modulate the relation between spatial language and spatial memory in deaf children and adults. *Memory & Cognition*, *51*(3), 582–600. doi:10.3758/s13421-022-01281-7.

Karadöller Astarlaroğlu, D. Z., Sümer, B., & Özyürek, A. (2017). Effects of delayed language exposure on spatial language acquisition by signing children and adults [Publication Type: Conference Paper / Full Text]. *Proceedings of the 39th Annual Meeting of the Cognitive Science Society: Computational Foundations of Cognition, CogSci 2017*, 2372–2376.

Karmiloff-Smith, A. (1979). Micro- and Macrodevelopmental Changes in Language Acquisition and Other Representational Systems [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1207/s15516709cog0302_1]. *Cognitive Science*, *3*(2), 91–118. doi:10.1207/s15516709cog0302_1.

Kayabaşı, D., & Gökgöz, K. (2023). The Causative-Inchoative Alternation and Age-of-Acquisition Effects on Multi-predicate Constructions in Turkish Sign Language. *Language Learning and Development*, *19*(2), 125–150. doi:10.1080/15475441.2022.2047690.

Kayabaşı, D., Sevgi, H., & Özsoy, A. S. (2020). Null arguments in Turkish Sign Language [Publisher: John Benjamins Publishing Company]. *Morphological Complexity within and across Boundaries: In honour of Asli Göksel*, *215*.

Kelepir, M. (2020). Morphology: 4. nominal inflection [Accessed 31-10-2021]. In M. Kelepir (Ed.), *A grammar of turkish sign language (tid)* (1st, pp. 230–237). https://thesignhub.eu/grammar/tid

Kelepir, M., Özkul, A., & Özparlak, E. T. (2018). Agent-backgrounding in turkish sign language (tid). *Sign Language & Linguistics*, *21*(2), 257–283.

Keleş, O. (2024). Applying computer vision to signed discourse: Effects of form and iconicity [Poster presented at the IMPRS Conference, Nijmegen].

Keleş, O., Atmaca, F., & Gökgöz, K. (2022). Effects of age of acquisition and category size on signed verbal fluency [Publisher: Routledge]. *Language Acquisition*, *29*(4), 361–383. doi:10.1080/10489223.2021.2023814.

Keleş, O., Atmaca, F., & Gökgöz, K. (2023). Reference tracking strategies of deaf adult signers in Turkish Sign Language [Publisher: Elsevier]. *Journal of Pragmatics*, *213*, 12–35.

Keleş, O., & Gökgöz, K. (2022). Expression of Aboutness Subject Topic Constructions in Turkish Sign Language (TİD) Narratives [Publisher: Edukacijsko-rehabilitacijski fakultet]. *Hrvatska revija za rehabilitacijska istraživanja*, *58*(Special Issue), 194–205. doi:10.31299/hrri.58.si.10.

Kemaloğlu, Y. K., & Kemaloğlu, P. Y. (2012). The history of sign language and deaf education in Turkey. *Kulak burun bogaz ihtisas dergisi: KBB = Journal of ear, nose, and throat*, *22*(2), 65–76. doi:10.5606/kbbihtisas.2012.013.

Kibrik, A., & Prozorova, E. V. (2007). Referential choice in signed and spoken languages. *DAARC 2007 (6th Discourse Anaphora and Anaphor Resolution Colloquium)*, 41–46.

Kimmelman, V., Imashev, A., Mukushev, M., & Sandygulova, A. (2020). Eyebrow position in grammatical and emotional expressions in Kazakh-Russian Sign Language: A quantitative study (M. L. Hall, Ed.). *PLOS ONE*, *15*(6), e0233731. doi:10.1371/journal.pone.0233731.

Kimmelman, V., Pfau, R., & Aboh, E. O. (2020). Argument structure of classifier predicates in Russian Sign Language. *Natural Language & Linguistic Theory*, *38*(2), 539–579. doi:10.1007/s11049-019-09448-9.

Kimmelman, V., & Teresė, A. (2023). Analyzing literary texts in Lithuanian Sign Language with Computer Vision: A proof of concept. In R. Galimullin & S. Touileb (Eds.), *CEUR Workshop Proceedings* (Vol. 3413).

Koulidobrova, E., & Lillo-Martin, D. (2016). A ‘point’of inquiry: The case of the (non-) pronominal IX in ASL. *The impact of pronominal form on interpretation*, *221*, 250.

Krebs, J., Malaia, E., Wilbur, R. B., Fessl, I., Wiesinger, H.-P., Schwameder, H., & Roehm, D. (2024, September). Motion Capture Analysis of Verb and Adjective Types in Austrian Sign Language [arXiv:2405.05161]. http://arxiv.org/abs/2405.05161

Kubuş, O. (2008). *An Analysis of Turkish Sign Language (TİD) Phonology and Morphology* [Master of Science]. Middle East Technical University.

Kuznetsova, A., Imashev, A., Mukushev, M., Sandygulova, A., & Kimmelman, V. (2021). Using computer vision to analyze non-manual marking of questions in KRSL. *Proceedings of the 1st International Workshop on Automatic Translation for Signed and Spoken Languages (AT4SSL)*, 49–59.

Leclercq, P., & Lenart, E. (2013). Discourse Cohesion and Accessibility of Referents in Oral Narratives: A Comparison of L1 and L2 Acquisition of French and English [Number: 12 Publisher: Presses universitaires de Caen]. *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics*, (12). doi:10.4000/discours.8801.

Lepic, R. (2019). A usage-based alternative to "lexicalization" in sign language linguistics [Number: 1 Publisher: Open Library of Humanities]. *Glossa: a journal of general linguistics*, *4*(1). doi:10.5334/gjgl.840.

Levy, E. T., & McNeill, D. (1992). Speech, gesture, and discourse. *Discourse Processes*, *15*(3), 277–301. doi:10.1080/01638539209544813.

Liddell, S. K. (2003). *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge University Press.

Lieberman, A. M., Borovsky, A., & Mayberry, R. I. (2018). Prediction in a visual language: Real-time sentence processing in American Sign Language across development. *Language, Cognition and Neuroscience*, *33*(4), 387–401. doi:10.1080/23273798.2017.1411961.

Lieberman, P. (1963). Some Effects of Semantic and Grammatical Context on the Production and Perception of Speech [Publisher: SAGE Publications Ltd]. *Language and Speech*, *6*(3), 172–187. doi:10.1177/002383096300600306.

Lillo-Martin, D., & Quadros, R. M. d. (2011). Acquisition of the syntax–discourse interface: The expression of point of view. *Lingua*, *121*(4), 623–636. doi:10.1016/j.lingua.2010.07.001.

Lin, S., Beddor, P. S., & Coetzee, A. W. (2014). Gestural reduction, lexical frequency, and sound change: A study of post-vocalic /l/. *Laboratory Phonology*, *5*(1). doi:10.1515/lp-2014-0002.

Lucas, C., Bayley, R., Rose, M., & Wulf, A. (2002). Location variation in american sign language. *Sign Language Studies*, 407–440.

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J., Chang, W.-T., Hua, W., Georg, M., & Grundmann, M. (2019, June). MediaPipe: A Framework for Building Perception Pipelines [arXiv:1906.08172 [cs]]. doi:10.48550/arXiv.1906.08172.

Makaroğlu, B. (2013). Interrogatives in turkish sign language (tid): The role of eyebrows. *Current directions in Turkish Sign Language research*, 143–165.

Malaia, E., & Wilbur, R. B. (2012). Kinematic Signatures of Telic and Atelic Events in ASL Predicates [Publisher: SAGE Publications Ltd]. *Language and Speech*, *55*(3), 407–421. doi:10.1177/0023830911422201.

Mandel, M. A. (1979). Natural constraints in sign language phonology: Data from anatomy [Publisher: JSTOR]. *Sign Language Studies*, (24), 215–229.

Marslen-Wilson, W., Levy, E., & Tyler, L. K. (1982). Producing interpretable discourse: The establishment and maintenance of reference [Publisher: Citeseer]. *Speech, place, and action*, 339–378.

Martínez-Guevara, N., & Curiel, A. (2024, May). Quantitative Analysis of Hand Locations in both Sign Language and Non-linguistic Gesture Videos. In E. Efthimiou, S.-E. Fotinea, T. Hanke, J. A. Hochgesang, J. Mesch, & M. Schulder (Eds.), *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources* (pp. 225–234). ELRA; ICCL. Retrieved October 14, 2024, from https://aclanthology.org/2024.signlang-1.25

Mauk, C. E., Lindblom, B., & Meier, R. P. (2008). Undershoot of ASL locations in fast signing. *Signs of the time. Selected papers from TISLR*, *8*, 3–24. Retrieved October 14, 2024, from http://www.signum-verlag.de/BTitel/pdf/978-3-936675-22-1BSP1.pdf

Mauk, C. E., & Tyrone, M. E. (2012). Location in ASL: Insights from phonetic variation. *Sign Language & Linguistics*, *15*(1), 128–146. doi:10.1075/sll.15.1.06mau.

Mauk, C. E. (2003). *Undershoot in two modalities: Evidence from fast speech and fast signing*. The University of Texas at Austin.

Mayberry, R. I. (2007). When timing is everything: Age of first-language acquisition effects on second-language learning [Publisher: Cambridge University Press]. *Applied Psycholinguistics*, *28*(3), 537–549. doi:10.1017/S0142716407070294.

McKee, R., Schembri, A., McKee, D., & Johnston, T. (2011). Variable "subject" presence in Australian Sign Language and New Zealand Sign Language. *Language Variation and Change*, *23*(3), 375–398. doi:10.1017/S0954394511000123.

Meier, R. P., & Newport, E. L. (1990). Out of the Hands of Babes: On a Possible Sign Advantage in Language Acquisition. *Language*, *66*(1), 1. doi:10.2307/415277.

Metzger, M. (1995). Constructed action and constructed dialogue in American Sign Language [Publisher: Gallaudet University Press Washington, DC]. *Sociolinguistics in Deaf communities*, 255–271.

Mitchell, R. E., & Karchmer, M. A. (2004). When Parents Are Deaf Versus Hard of Hearing: Patterns of Sign Use and School Placement of Deaf and Hard-of-Hearing Children [Publisher: Oxford Academic]. *The Journal of Deaf Studies and Deaf Education*, *9*(2), 133–152. doi:10.1093/deafed/enh017.

Moon, S.-J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical Society of America*, *96*(1), 40–55. doi:10.1121/1.410492.

Morgan, G. (2000). Discourse cohesion in sign and speech. *International Journal of Bilingualism*, *4*(3), 279–300. doi:10.1177/13670069000040030101.

Morgan, G. (2005). Advances in the Sign-Language Development of Deaf Children [Publication Title: Advances in the Sign-Language Development of Deaf Children Section: Advances in the Sign-Language Development of Deaf Children]. In *The Development of Narrative Skills in British Sign Language*. Oxford University Press.

Napoli, D. J., Sanders, N., & Wright, R. (2014). On the Linguistic Effects of Articulatory Ease, with a Focus on Sign Languages [Publisher: Linguistic Society of America]. *Language*, *90*(2), 424–456. Retrieved October 13, 2024, from https://www.jstor.org/stable/24672181

Needham, L., Evans, M., Cosker, D. P., Wade, L., McGuigan, P. M., Bilzon, J. L., & Colyer, S. L. (2021). The accuracy of several pose estimation methods for 3D joint centre localisation [Publisher: Nature Publishing Group UK London]. *Scientific reports*, *11*(1), 20673.

Nespor, M., & Sandler, W. (1999). Prosody in Israeli sign language [Publisher: SAGE Publications Sage UK: London, England]. *Language and speech*, *42*(2-3), 143–176.

Newport, E. L. (1990). Maturational Constraints on Language Learning [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1207/s15516709cog1401_2]. *Cognitive Science*, *14*(1), 11–28. doi:10.1207/s15516709cog1401_2.

Normoyle, A., Artacho, B., Savakis, A., Senghas, A., Badler, N. I., Occhino, C., Rothstein, S. J., & Dye, M. W. G. (2022). Open-source pipeline for skeletal modeling of sign language utterances from 2d video sources. *Theoretical Issues in Sign Language Research (TISLR)*.

Nuhbalaoglu, D. (2018). *Comprehension and production of referential expressions in German Sign Language and Turkish Sign Language: An empirical approach* [Doctoral Dissertation]. Georg-August-Universität Göttingen.

Ormel, E., Crasborn, O., Kootstra, G. J., & De Meijer, A. (2017). Coarticulation of Handshape in Sign Language of the Netherlands: A Corpus Study. *Laboratory Phonology*, *8*(1), 10. doi:10.5334/labphon.45.

Östling, R., Börstell, C., & Courtaux, S. (2018). Visual iconicity across sign languages: Large-scale automated video analysis of iconic articulators and locations [Publisher: Frontiers Media SA]. *Frontiers in psychology*, *9*, 725.

Özsoy, S., & Nuhbalaoğlu, D. (2014). Linearization in noun phrases in Turkish Sign Language. *Formal and experimental approaches in sign languages*, 9–11.

Özyürek, A., Furman, R., & Goldin-Meadow, S. (2015). On the way to language: Event segmentation in homesign and gesture. *Journal of Child Language*, *42*(1), 64–94. doi:10.1017/S0305000913000512.

Padden, C. (1986). Verbs and role-shifting in American Sign Language. *Proceedings of the fourth national symposium on sign language research and teaching*, *44*, 57.

Perniss, P., & Özyürek, A. (2015). Visible Cohesion: A Comparison of Reference Tracking in Sign, Speech, and Co-Speech Gesture [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/tops.12122]. *Topics in Cognitive Science*, *7*(1), 36–60. doi:10.1111/tops.12122.

Perniss, P., Pfau, R., & Steinbach, M. (2007, August). Can't you see the difference? Sources of variation in sign language structure. In P. M. Perniss, R. Pfau, & M. Steinbach (Eds.), *Visible Variation* (pp. 1–34). Mouton de Gruyter. doi:10.1515/9783110198850.1.

Pfau, R. (2011). A point well taken. *Deaf around the world: The impact of language*, 144–163.

Pizzuto, E., Rossini, P., Sallandre, M.-A., & Wilkinson, E. (2006). Deixis, Anaphora and Highly Iconic Structures: Cross-linguistic Evidence on American (ASL), French (LSF) and Italian (LIS) Signed Languages. *TISLR9*, 475–495. https://hal.science/hal-00617907

Prince, E. F. (1978). A comparison of wh-clefts and it-clefts in discourse [Publisher: JSTOR]. *Language*, 883–906.

Puupponen, A., Kanto, L., Wainio, T., & Jantunen, T. (2022). Variation in the use of constructed action according to discourse type and age in Finnish Sign Language [Publisher: Elsevier]. *Language & Communication*, *83*, 16–35.

Puupponen, A., Wainio, T., Burger, B., & Jantunen, T. (2015). Head movements in Finnish Sign Language on the basis of Motion Capture data: A study of the form and function of nods, nodding, head thrusts, and head pulls [Publisher: John Benjamins]. *Sign Language & Linguistics*, *18*(1), 41–89. doi:10.1075/sll.18.1.02puu.

Pyers, J. E., & Senghas, A. (2009). Language Promotes False-Belief Understanding: Evidence From Learners of a New Sign Language. *Psychological Science*, *20*(7), 805–812. doi:10.1111/j.1467-9280.2009.02377.x.

Ramírez, N. F., Lieberman, A. M., & Mayberry, R. I. (2013). The initial stages of first-language acquisition begun in adolescence: When late looks early. *Journal of Child Language*, *40*(2), 391–414. doi:10.1017/S0305000911000535.

Ranum, O., Otterspeer, G., Andersen, J. I., Belleman, R. G., & Roelofsen, F. (2024, May). 3D-LEX v1.0 – 3D Lexicons for American Sign Language and Sign Language of the Netherlands. In E. Efthimiou, S.-E. Fotinea, T. Hanke, J. A. Hochgesang, J. Mesch, & M. Schulder (Eds.), *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources* (pp. 290–301). ELRA; ICCL. Retrieved October 14, 2024, from https://aclanthology.org/2024.signlang-1.33

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints [Publisher: Acoustical Society of America]. *The Journal of the Acoustical Society of America*, *102*(1), 544–561.

Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics [Publisher: Department of Philosophy and Moral Science at Ghent University.]. *Philosophica*, *27*.

Roh, K., Lee, H., Hwang, E. J., Cho, S., & Park, J. C. (2024). Preprocessing Mediapipe Keypoints with Keypoint Reconstruction and Anchors for Isolated Sign Language Recognition. *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources*, 323–334. Retrieved October 14, 2024, from https://aclanthology.org/2024.signlang-1.36.pdf

Russell, K., Wilkinson, E., & Janzen, T. (2011). ASL sign lowering as undershoot: A corpus study [Publisher: De Gruyter Mouton]. *Laboratory Phonology*, *2*(2), 403–422. doi:10.1515/labphon.2011.015.

Ryan, J. (2015). Overexplicit referent tracking in l2 english: Strategy, avoidance, or myth? *Language Learning*, *65*(4), 824–859.

Ryan, M.-L. (2003). On defining narrative media. *Image & Narrative*, *6*, 1–7.

Sáenz, M. D. C. (2022). Mouthing Recognition with OpenPose in Sign Language. *Proceedings of the 7th International Workshop on Sign Language Translation and Avatar Technology: The Junction of the Visual and the Textual: Challenges and Perspectives*, 91–94. Retrieved October 14, 2024, from https://aclanthology.org/2022.sltat-1.14/

Şahin, K., & Gökgöz, K. (2024, May). Decoding Sign Languages: The SL-FE Framework for Phonological Analysis and Automated Annotation. In E. Efthimiou, S.-E. Fotinea, T. Hanke, J. A. Hochgesang, J. Mesch, & M. Schulder (Eds.), *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign*

Language Resources (pp. 335–342). ELRA; ICCL. Retrieved October 14, 2024, from https://aclanthology.org/2024.signlang-1.37

Sanders, N., & Napoli, D. J. (2016). Reactive effort as a factor that shapes sign language lexicons. *Language*, *92*(2), 275–297. doi:10.1353/lan.2016.0032.

Sandler, W. (1999). Cliticization and prosodic words in a sign language. *Studies on the phonological word*, *174*, 223.

Sari, H. (2005). An Analysis of the Relationship Between Identity Patterns of Turkish Deaf Adolescents and the Communication Modes Used in Special Residential Schools for the Hearing Impaired and Deaf. *Deafness & Education International*, *7*(4), 206–222. doi:10.1179/146431505790560329.

Schembri, A. (2003). Rethinking 'classifiers' in signed languages. In *Perspectives on classifier constructions in sign languages* (pp. 13–44). Psychology Press.

Schlenker, P. (2017). Sign Language and the Foundations of Anaphora. *Annual Review of Linguistics*, *3*(1), 149–177. doi:10.1146/annurev-linguistics-011415-040715.

Selting, M. (2007). Lists as embedded structures and the prosody of list construction as an interactional resource. *Journal of Pragmatics*, *39*(3), 483–526.

Senghas, A., & Coppola, M. (2001). Children Creating Language: How Nicaraguan Sign Language Acquired a Spatial Grammar [Publisher: SAGE Publications Inc]. *Psychological Science*, *12*(4), 323–328. doi:10.1111/1467-9280.00359.

Senghas, A., Kita, S., & Özyürek, A. (2004). Children Creating Core Properties of Language: Evidence from an Emerging Sign Language in Nicaragua [Publisher: American Association for the Advancement of Science Section: Report]. *Science*, *305*(5691), 1779–1782. doi:10.1126/science.1100199.

Sevgi, H., & Gökgöz, K. (2023). Classifiers, argument expression, and age of acquisition effects in turkish sign language (tid). *Sign Language & Linguistics*, *26*(1), 1–36.

Sevinç, A. M. (2006). Grammatical relations and word order in Turkish Sign Language (TİD). *Unpublished master's thesis, Middle East Technical University, Turkey*.

Sloetjes, H., & Wittenburg, P. (2008). Annotation by Category: ELAN and ISO DCR. *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*.

Slonimska, A., Di Renzo, A., Kanakanti, M., Campisi, E., Capirci, O., & Ozyurek, A. (2024). Iconicity as a communicative strategy in sign languages: Quantitative vs. qualitative modulations of constructed action for children and adults in LIS (Italian Sign Language).

Smith, S., & Cormier, K. (2014). In or Out?: Spatial Scale and Enactment in Narratives of Native and Nonnative Signing Deaf Children Acquiring British Sign Language. *Sign Language Studies*, *14*(3), 275–301. doi:10.1353/sls.2014.0008.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism [Publisher: John Benjamins]. *Linguistic approaches to bilingualism*, *1*(1), 1–33.

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of italian. *Second language research*, *22*(3), 339–368.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition* (Vol. 142). Harvard University Press Cambridge, MA.

Stamp, R., Cohn, D., Hel-Or, H., & Sandler, W. (2024). Kinect-ing the Dots: Using Motion-Capture Technology to Distinguish Sign Language Linguistic From Gestural Expressions. *Language and Speech*, *67*(1), 255–276. doi:10.1177/00238309231169502.

Stamp, R., Dachkovsky, S., Hel-Or, H., Cohe, D., & Sandler, W. (2024). A kinematic study of phonetic reduction in a young sign language. *Journal of Phonetics*, *104*, 101311. doi:10.1016/j.wocn.2024.101311.

Stewart, J. (2014). A quantitative analysis of sign lengthening in american sign language. *Sign Language & Linguistics*, *17*(1), 82–101.

Supalla, T. (1986). The classifier system in American sign language. *Noun classes and categorization*, *7*, 181–214.

Swabey, L. A. (2002). *The Cognitive Status, Form and Distribution of Referring Expressions in ASL and English Narratives* [Unpublished Doctoral Dissertation]. University of Minnesota.

Talmy, L. (2003). The representation of spatial structure in spoken and signed language [Publisher: Mahwah, New Jersey: Lawrence Erlbaum Associates]. *Perspectives on classifier constructions in sign language*, 169–195.

Taşçı, S., & Göksel, A. (2018). Native compounds in tid: A classification based on headedness. *Dilbilim Araştırmaları Dergisi*, *29*(1), 1–28.

Tily, H., & Piantadosi, S. (2009). Refer efficiently: Use less informative expressions for more predictable meanings. *Proceedings of the workshop on the production of referring expressions: Bridging the gap between computational and empirical approaches to reference*, 1–8.

Tomlin, R. S. (1987). Linguistic reflections of cognitive events [Publisher: John Benjamins Amsterdam]. *Coherence and grounding in discourse*, *11*, 455–479.

Toole, J. (1996, June). The Effect of Genre on Referential Choice. In T. Fretheim & J. K. Gundel (Eds.), *Reference and Referent Accessibility* (p. 263). John Benjamins Publishing Company. doi:10.1075/pbns.38.16too.

Türkiye. (2005). Disability law (under the disability act subsection no. 5378) [Enacted in 2005 under Turkish Disability Act].

Tyrone, M. E., & Mauk, C. E. (2010). Sign lowering and phonetic reduction in American Sign Language. *Journal of Phonetics*, *38*(2), 317–328. doi:10.1016/j.wocn.2010.02.003.

Tyrone, M. E., & Mauk, C. E. (2012). Phonetic reduction and variation in American Sign Language: A quantitative study of sign lowering. *Laboratory Phonology*, *3*(2). doi:10.1515/lp-2012-0019.

Van der Hulst, H. (1993). Units in the analysis of signs. *Phonology*, *10*(2), 209–241.

Walker, M., Joshi, A., & Prince, E. (1997, December). Centering in Naturally Occurring Discourse: An Overview. In M. A. Walker, A. K. Joshi, & E. F. Prince (Eds.), *Centering Theory In Discourse* (pp. 1–28). Oxford University Press. doi:10.1093/oso/9780198236870.003.0001.

Wang, T. (2022). *Examining the Role of Subjecthood and Pronominalization in Reference Resolution in L1 and L2 Processing: An Eye-Tracking Study* [PhD Thesis]. University of Kansas.

Wilbur, R. B., Malaia, E., & Shay, R. A. (2012). Degree modification and intensification in american sign language adjectives. *Logic, Language and Meaning: 18th Amsterdam Colloquium, Amsterdam, The Netherlands, December 19-21, 2011, Revised Selected Papers*, 92–101.

Williams, J. (1988). Zero Anaphora in Second Language Acquisition: A Comparison among Three Varieties of English [Publisher: Cambridge University Press]. *Studies in Second Language Acquisition*, *10*(3), 339–370. doi:10.1017/S0272263100007488.

Woll, B., Meurant, L., Sinte, A., Van Herreweghe, M., & Vermeerbergen, M. (2013). Sign language and spoken language development in young children:: Measuring vocabulary by means of the CDI. In *Sign Language Research, Uses and Practices* (1st ed., pp. 15–34). De Gruyter. https://www.jstor.org/stable/j.ctvbkk4dr.4

Wulf, A., Dudis, P., Bayley, R., & Lucas, C. (2002). Variable subject presence in ASL narratives [Publisher: JSTOR]. *Sign Language Studies*, 54–76.

Yaşar, A., Kisbet, B., Şahin, K., & Gökgöz, K. (2024). Complexity of telicity marking in Turkish Sign Language (TİD) [Poster presented at the IMPRS Conference, Nijmegen].

Zeshan, U. (2003). Aspects of Türk Isaret Dili (Turkish Sign Language) [Publisher: John Benjamins]. *Sign Language & Linguistics*, *6*(1), 43–75. doi:10.1075/sll.6.1.04zes.

Zhang, Z. (2012). Microsoft Kinect Sensor and Its Effect [Conference Name: IEEE MultiMedia]. *IEEE MultiMedia*, *19*(2), 4–10. doi:10.1109/MMUL.2012.24.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley.

Zwitserlood, I. (2012). Classifiers. In *Sign language: An international handbook*. De Gruyter.